

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:34:08 ; Search time 2172.55 Seconds

(without alignments)  
10796.945 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806  
Sequence: 1 cgcgcacagctcgggacaa.....aagagagagagagagacta 806

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenBank.\*  
1: gb.ba.\*  
2: gb.hg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vl.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vl.\*  
30: em.hg.hum.\*  
31: em.hg.in.\*  
32: em.hg.other.\*  
33: em.hg.mus.\*  
34: em.hg.pln.\*  
35: em.hg.rtd.\*  
36: em.hg.mam.\*  
37: em.hg.vrt.\*  
38: em.sy.\*  
39: em.hggo.hum.\*  
40: em.hggo.mus.\*  
41: em.hggo.other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	750.4	93.1	814	6	AX440456
2	750.4	93.1	814	9	AB015631
3	749.4	93.0	1210	6	AX464040
4	743.4	92.2	1080	9	AY032624
5	590.4	73.3	158198	10	BC008261
6	583.2	72.4	1017	10	AF186115
7	583.2	72.4	1017	10	AF186115
8	308.6	38.3	649	9	BC001027
9	288.4	35.8	824	9	AX440375
10	237.8	23.5	564	6	AC025574
11	206	25.6	182350	2	AC073896
12	206	25.6	182350	2	AC023500
13	206	25.6	182350	2	AC023500
14	177.2	22.0	207424	10	AC090489
15	174	21.6	101566	2	AC109891
16	147	18.2	147	6	AX333297
17	147	18.2	147	6	AX408688
18	126	15.6	126	6	AX333296
19	126	15.6	126	6	AX408687
20	112	13.9	69753	2	AC103156
21	94	11.7	95437	2	AC096788
22	84.4	10.5	2378	9	AK094445
23	56	6.9	622	5	AB055671
24	55.4	6.9	484	9	HSW803189
25	53.4	6.6	996	3	AY094698
26	51.4	6.4	55359	3	AC012854
27	51.4	6.4	172904	3	AC007414
28	51.4	6.4	273590	3	AE003831
29	48.6	6.0	3522	9	HSW804706
30	48.4	6.0	3833	9	AF399910
31	48.4	6.0	12393	6	AX346165
32	47.8	5.9	168994	9	AC097480
33	47.6	5.9	2640	9	AF184965
34	47.2	5.9	1534	9	HSW801279
35	47.2	5.9	1929	10	NM005931A
36	47.2	5.9	2446	9	HSW805337
37	47	5.8	168418	9	AC009403
38	46.8	5.8	94871	2	AC111353
39	46.8	5.8	129360	2	AC117079
40	46.4	5.8	1053	5	AF460207
41	46.4	5.8	3499	10	BC007476
42	46.4	5.8	183982	2	AC009825
43	46.2	5.7	63365	2	AC087437
44	46.2	5.7	94804	2	AL358935
45	46.2	5.7	144984	9	HS4333F14

ALIGNMENTS

RESULT 1  
LOCUS AX440456 814 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 309 from Patent WO0190154.  
ACCESSION AX440456  
VERSION AX440456.1 GI:21665266  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secret, H.,  
Lodes, M.J., Algate, P.A., Fling, S.P., Mannion, D., Benson, D.R. and  
Carter, D.

TITLE Compositions and methods for the therapy and diagnosis of ovarian cancer  
 JOURNAL Patent: WO 0190154-A 309 29-NOV-2001;  
 CORIXA CORPORATION (US)  
 FEATURES Location/Qualifiers  
 source 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Query Match 93.1%; Score 750.4; DB 6; Length 814;  
 Best Local Similarity 99.0%; Pred. No. 4.9e-183;  
 Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

OY 11 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCCGCTTGGGCGCAGCGGGCGCCCTGG 70
DB 52 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCCGCTTGGGCGCAGCGGGCGCCCTGG 111
OY 71 GACCAAGGTGAGCAACCCCGTTACCCCTAARATGAAGAGCTGGGCTTGCTGGCCCTG 130
DB 112 GACGAGAGTGGAGCGAGCCCATTAAGCTAAGATGAAGAGCTGGGCTTGCTGGCCCTG 171
OY 131 CTTTGGGGGGCCCTGCTGGAGACCCGCTGGGCTGCGAGAGAGCAGAGATCTCACTGTGA 190
DB 172 CTTTGGGGGGCCCTGCTGGAGACCCGCTGGGCTGCGAGAGAGCAGAGATCTCACTGTGA 231
OY 191 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 250
DB 232 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 291
OY 251 ACCATTACAGATGGATCTTCCGGATCAATCCAGATGGACCAAGTCACTGTGAGGTG 310
DB 292 ACCATTACAGATGGATCTTCCGGATCAATCCAGATGGACCAAGTCACTGTGAGGTG 351
OY 311 CCTTATGCCCCGCTCAGAGGCCCCACCTCACAGAGTGTGTGAGAGATATGACCGGATG 370
DB 352 CCTTATGCCCCGCTCAGAGGCCCCACCTCACAGAGTGTGTGAGAGATATGACCGGATG 411
OY 371 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 430
DB 412 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 471
OY 431 GCGCGGAATGAGATCCAGTCACTGACCTACAGAGGATCCGATCCGATCAGATATT 490
DB 472 GCGCGGAATGAGATCCAGTCACTGACCTACAGAGGATCCGATCCGATCAGATATT 531
OY 491 ACGGCAACCTCAAGTTTCTGTGTGAGACATTTGTGAGAGATPACGAGATCACTATT 550
DB 532 ACGGCAACCTCAAGTTTCTGTGTGAGACATTTGTGAGAGATPACGAGATCACTATT 591
OY 551 GAATCTTTTCCGAGAGGCTGACATGTAAAGACAACTTTCAGTAAAGCAACAT 610
DB 592 GAATCTTTTCCGAGAGGCTGACATGTAAAGACAACTTTCAGTAAAGCAACAT 651
OY 611 CTTTGTGACCAATGCCCTGACATATGCGATGAGCTATGAACCACTGGACAGCCCAAC 670
DB 652 CTTTGTGACCAATGCCCTGACATATGCGATGAGCTATGAACCACTGGACAGCCCAAC 711
OY 671 ACTGGCTTATGATGATCAACCCCGAGAGGAGAAATGGTGGCAATGCTTTATATTAT 730
DB 712 ACTGGCTTATGATGATCAACCCCGAGAGGAGAAATGGTGGCAATGCTTTATATTAT 771
OY 731 GTTTTACTGAATTAAGTGAATAAATATGAACCAAAAGTA 772
DB 772 GTTTTACTGAATTAAGTGAATAAATATGAACCAAAAGTA 813

```

RESULT 2  
 AB015631  
 LOCUS AB015631 814 bp mRNA linear PRI 26-JUL-2001  
 DEFINITION Homo sapiens mRNA for type II membrane protein, complete cds,  
 clone:HP10390.

ACCESSION AB015631  
 VERSION AB015631.1 GI:4586639  
 KEYWORDS type II membrane protein.  
 SOURCE Homo sapiens gastric adenocarcinoma cDNA to mRNA,  
 clone.lib:pkal-meta-1 clone:HP10390.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1  
 Yokoyama-Kobayashi,M., Yamaguchi,T., Sekine,S. and Kato,S.  
 Selection of cDNAs encoding putative type II membrane proteins on  
 the cell surface from a human full-length cDNA bank  
 Gene 228 (1-2), 161-167 (1999)  
 2 (bases 1 to 814)  
 Kato,S.  
 Direct Submission  
 Submitted (16-JUN-1998) Selski Kato, Research Institute of National  
 Rehabilitation Engineering; 4-1 Namiki, Tokorozawa, Saitama  
 359-8555, Japan (E-mail:selski@rehab.go.jp,  
 Tel:042-995-3100(ex.2568), Fax:042-995-3132)  
 FEATURES Location/Qualifiers  
 source 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HP10390"  
 /issue\_type="gastric adenocarcinoma"  
 /clone.lib="pkal-meta-1"  
 145..693  
 /codon\_start=1  
 /product="type II membrane protein"  
 /protein\_id="BAI76498.1"  
 /db\_xref="GI:4586640"  
 /translation="MKGSGWLLALLGALLGTAMARRSODLHCAGCRALVDELEVEIAD  
 VDKPTOMGSEFRINPQSSQSVVPPARSAHLTELEELCEIDMKKEGQIDSTHR  
 KNYRVYGRNESESLDQGRIDSDISGTLKFPACESIYEIEDELLIFPSREADNYK  
 DKLSKRDLCDHALHSHDEL"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Query Match 93.1%; Score 750.4; DB 9; Length 814;  
 Best Local Similarity 99.0%; Pred. No. 4.9e-183;  
 Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

OY 11 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCCGCTTGGGCGCAGCGGGCGCCCTGG 70
DB 52 CTGGGGCCCAAGTGAAGATCCAGCGCTGCGCCGCTTGGGCGCAGCGGGCGCCCTGG 111
OY 71 GACCAAGGTGAGCAACCCCGTTACCCCTAARATGAAGAGCTGGGCTTGCTGGCCCTG 130
DB 112 GACGAGAGTGGAGCGAGCCCATTAAGCTAAGATGAAGAGCTGGGCTTGCTGGCCCTG 171
OY 131 CTTTGGGGGGCCCTGCTGGAGACCCGCTGGGCTGCGAGAGAGCAGAGATCTCACTGTGA 190
DB 172 CTTTGGGGGGCCCTGCTGGAGACCCGCTGGGCTGCGAGAGAGCAGAGATCTCACTGTGA 231
OY 191 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 250
DB 232 GCATGAGAGGCTCTGTGTGATGACATAGAAATGGAAATGGCCAGGTGAGCCCAAGAAG 291
OY 251 ACCATTACAGATGGATCTTCCGGATCAATCCAGATGGACCAAGTCACTGTGAGGTG 310
DB 292 ACCATTACAGATGGATCTTCCGGATCAATCCAGATGGACCAAGTCACTGTGAGGTG 351
OY 311 CCTTATGCCCCGCTCAGAGGCCCCACCTCACAGAGTGTGTGAGAGATATGACCGGATG 370
DB 352 CCTTATGCCCCGCTCAGAGGCCCCACCTCACAGAGTGTGTGAGAGATATGACCGGATG 411
OY 371 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 430
DB 412 AAGGATATGGGGAACAGATTGATCTTCCACCATCCGAGAGACTAGCTAGTGTAGTG 471

```



QY 431 GGCAGGATGAGATTCAGTGAACCTGACCTACAGGATCCGATCGATCATATT 490  
|||||  
Db 472 GGCAGGATGAGATTCAGTGAACCTGACCTACAGGATCCGATCGATCATATT 531  
QY 491 AGGGGACCTCCAGTGTGGTGAGAGATTTGGAGGAATACGAGATGAATCATT 550  
|||||  
Db 532 AGGGGACCTCCAGTGTGGTGAGAGATTTGGAGGAATACGAGATGAATCATT 591  
QY 551 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAAGCAAGAT 610  
|||||  
Db 592 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAAGCAAGAT 651  
QY 611 CTTTGTGACCATGACCTGACATATGCGATGATGACGTATGAACCACTGAGAGCCAC 670  
|||||  
Db 652 CTTTGTGACCATGACCTGACATATGCGATGATGACGTATGAACCACTGAGAGCCAC 711  
QY 671 ACTGCTTGTGATGATCACCCTGACAGAGGGAATATGTTGCAATGCTTTATATATAT 730  
|||||  
Db 712 ACTGCTTGTGATGATCACCCTGACAGAGGGAATATGTTGCAATGCTTTATATATAT 771  
QY 731 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGTA 772  
|||||  
Db 772 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGTA 813

RESULT 3  
AX464040  
LOCUS AX464040 1210 bp DNA linear PAT 16-JUL-2002  
DEFINITION Sequence 173 from Patent WO0140466.  
ACCESSION AX464040  
VERSION AX464040.1 GI:21899037  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS 1  
Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,  
Geo,W.O., Gerlielsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,  
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,  
Wood,W.L. and Zhang,Z.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding  
same  
JOURNAL Patent: WO 0140466-A 173 07-JUN-2001;  
Genentech Inc. (US)  
FEATURES  
source 1..1210  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 261 a 316 c 379 g 254 t  
ORIGIN

Query Match 93.0%; Score 749.4; DB 6; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 9.6e-183;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 11 CTTGGGCGCAAGTGAAGATCCAGCGCTGCGACGCGCTTGGGCGCGCGCCCTGG 70  
|||||  
Db 450 CTTGGGCGCAAGTGAAGATCCAGCGCTGCGACGCGCTTGGGCGCGCGCCCTGG 509  
QY 71 GACCAAAAGGTGAGCAACCCGCTTACCTAAARATGAAGGCTGGGCTGGCGCCG 130  
|||||  
Db 510 GACCAAAAGGTGAGCAACCCGCTTACCTAAARATGAAGGCTGGGCTGGCGCCG 569  
QY 131 CTTTGTGAGGCGCTGCTGCTGGAACCCGCTGCGAGAGAGCGAGATCTCACTGTGA 190  
|||||  
Db 570 CTTTGTGAGGCGCTGCTGCTGGAACCCGCTGCGAGAGAGCGAGATCTCACTGTGA 629  
QY 191 GCATGAGGCGCTGCTGCTGGAATAGATGGAATGGGAAATTTGCCAGGTGAGACCCCAAGA 250  
|||||  
Db 630 GCATGAGGCGCTGCTGCTGGAATAGATGGAATGGGAAATTTGCCAGGTGAGACCCCAAGA 689  
QY 251 ACCATTCAATGGATCTTTCCGGAATCAATCAAGATGGACGACGATGCTGGTGAAGT 310

Db 690 ACCATTCAATGGATCTTTCCGGAATCAATCAAGATGGACGACGATGCTGGTGAAGT 749  
QY 311 CTTTATGCGCGCTCAGAGGCCACCTCAGACAGCTGCTGAGAGAGATATGTCAGCGATG 370  
|||||  
Db 750 CTTTATGCGCGCTCAGAGGCCACCTCAGACAGCTGCTGAGAGAGATATGTCAGCGATG 809  
QY 371 AAGAGATGAGGAGACAGATGATGCTTCCACCATCGGAAGAACTACGTACTGTAGT 430  
|||||  
Db 810 AAGAGATGAGGAGACAGATGATGCTTCCACCATCGGAAGAACTACGTACTGTAGT 869  
QY 431 GGCAGGATGAGATTCAGTGAACCTGACCTCAAGAGCATCGATCGATCATATT 490  
|||||  
Db 870 GGCAGGATGAGATTCAGTGAACCTGACCTCAAGAGCATCGATCGATCATATT 929  
QY 491 AGGGGACCTCCAGTGTGGTGAGAGATTTGGAGGAATACGAGATGAATCATT 550  
|||||  
Db 930 AGGGGACCTCCAGTGTGGTGAGAGATTTGGAGGAATACGAGATGAATCATT 989  
QY 551 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAAGCAAGAT 610  
|||||  
Db 990 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAAGCAAGAT 1049  
QY 611 CTTTGTGACCATGACCTGACATATGCGATGATGACGTATGAACCACTGAGAGCCAC 670  
|||||  
Db 1050 CTTTGTGACCATGACCTGACATATGCGATGATGACGTATGAACCACTGAGAGCCAC 1109  
QY 671 ACTGCTTGTGATGATCACCCTGACAGAGGGAATATGTTGCAATGCTTTATATATAT 730  
|||||  
Db 1110 ACTGCTTGTGATGATCACCCTGACAGAGGGAATATGTTGCAATGCTTTATATATAT 1169  
QY 731 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGT 771  
|||||  
Db 1170 GTTTTACTGAAATTAATGAAAAAATATGAAACCAAAAGT 1210

RESULT 4  
AY032624  
LOCUS AY032624 1080 bp mRNA linear PRI 18-APR-2002  
DEFINITION Homo sapiens saposin-like protein mRNA, complete cds.  
ACCESSION AY032624  
VERSION AY032624.1 GI:20196198  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS 1 (bases 1 to 1080)  
Bornhauser,B.C., Olsson,P.-A. and Lindholm,D.  
TITLE NSAP is a novel saposin-like protein that interacts with MIR and  
stimulates neurite outgrowth  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1080)  
Olsson,P.-A. and Lindholm,D.  
AUTHORS Direct Submission  
TITLE Submitted (13-APR-2001) Neuroscience, Uppsala University,  
Husargatan 3, Uppsala 75123, Sweden  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..1080  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q15"  
419..967  
/note="NSAP: encodes a type 2 membrane signal and a  
putative ER retention motif."  
/codon\_start=1  
/product="saposin-like protein"  
/protein\_id="AAK3148.1"  
/db\_xref="GI:20196198"  
/translation="MKGMWALLLGLALGLTAAARSRSDILGCAALVDELEINQ  
VDPKTIOMGSFRINPDGSSVVEPIPAKRSFAHLTLEELICDRKKEGQIDPSTHR  
KNYRVVGRNSESSELDGIRIDISIGTLKFACEISIVEYEDELIEFFSREADNVK

BASE COUNT 238 a 269 c 343 g 230 t  
 ORIGIN  
 Query Match 92.2%; Score 743.4; DB 9; Length 1080;  
 Best Local Similarity 98.9%; Pred. No. 3.3e-181;  
 Matches 747; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCCAAGTGAAGTCCAGCGGCTCTCCAGCGCTTGGGCGACGGGGGCGCTGG 70  
 326 CTGGGGCCCAAGTGAAGTCCAGCGGCTCTCCAGCGCTTGGGCGACGGGGGCGCTGG 385  
 71 GACCAAGTGGAGCAACCCCGTTACCTAAATGAAGAGTGGGCTGGGCTGGGCTGG 130  
 386 GAGCAAGTGGAGCAACCCCGTTACCTAAATGAAGAGTGGGCTGGGCTGGGCTGG 445  
 131 CTCTGGGGGCGCTCTGGGAGCAACCGGCTGGGCTGGGAGGAGGAGGATCTCCACTGGA 190  
 446 CTCTGGGGGCGCTCTGGGAGCAACCGGCTGGGCTGGGAGGAGGAGGATCTCCACTGGA 505  
 191 GCATGAGGCGCTCTGGTGGATGAATGAATGGAAATGGCCAGGTGACCCCAAGAG 250  
 506 GCATGAGGCGCTCTGGTGGATGAATGAATGGAAATGGCCAGGTGACCCCAAGAG 565  
 251 ACATTCAGATGGATCTTCCGGATCAATCCAGATGGGAGGAGGAGGAGGAGGAG 310  
 566 ACATTCAGATGGATCTTCCGGATCAATCCAGATGGGAGGAGGAGGAGGAGGAG 625  
 311 CCTATGCGCGCTCTGAGAGGCGCCAGCTCAGAGAGCTCTGGAGAGATATGTACCGGATG 370  
 626 CCTATGCGCGCTCTGAGAGGCGCCAGCTCAGAGAGCTCTGGAGAGATATGTACCGGATG 685  
 371 AAGAGATATGGGAGACAGATGTATCTCCAGGATCGGAGAGATATGTACCGGATG 430  
 686 AAGAGATATGGGAGACAGATGTATCTCCAGGATCGGAGAGATATGTACCGGATG 745  
 431 GCGCGGAATGGAGAACCCAGTCAATCCAGATGGGAGGAGGAGGAGGAGGAGGAG 490  
 746 GCGCGGAATGGAGAACCCAGTCAATCCAGATGGGAGGAGGAGGAGGAGGAGGAG 805  
 491 AGCGGACCCCTCAAGTTCGCTGAGAGCATTGTGGAGAGATACGAGATGAATCAAT 550  
 806 AGCGGACCCCTCAAGTTCGCTGAGAGCATTGTGGAGAGATACGAGATGAATCAAT 865  
 551 GAATTCCTTCCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACAGTAAAGACAGAT 610  
 866 GAATTCCTTCCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACAGTAAAGACAGAT 925  
 611 CTTTGGACATCCCTCGACATATGCAATGAGTGAACCACTGAGAGGAGGAGGAG 670  
 926 CTTTGGACATCCCTCGACATATGCAATGAGTGAACCACTGAGAGGAGGAGGAG 985  
 671 ACTGCGTATGATGATCAACCCAGAGAGGAGAAATGTTGAGATGCTTTATATATAT 730  
 986 ACTGCGTATGATGATCAACCCAGAGAGGAGAAATGTTGAGATGCTTTATATATAT 1045  
 731 GTTTTACTGAATTAATGAAAAATATGAAC 765  
 1046 GTTTTACTGAATTAATGAAAAATATGAAC 1080

RESULT 5  
 BC008261 798 bp mRNA linear ROD 07-AUG-2002  
 LOCUS  
 DEFINITION Mus musculus, transmembrane protein 4, clone MGC:6853  
 IMAGE:2650612, mRNA, complete cds.  
 ACCESSION BC008261  
 VERSION BC008261.1 GI:14198400  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 798)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hui, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

FEATURES  
 source  
 1..798  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="129.C57BL/6J,FVB/N"  
 /clone="MGC:6853 IMAGE:2650612"  
 /tissue="type="Mammary tumor, Brca1-/fl; MMTV-Cre model. 10 months old, gross tissue."  
 /clone\_id="NCI CGAP\_Mam3"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-SPORT6"  
 115..663  
 /codon\_start=1  
 /product="transmembrane protein 4"  
 /protein\_id="AA08261.1"  
 /db\_xref="GI:14198401"  
 /db\_xref="locusid:56530"  
 /translation="MKGWMLALLGLVLTGTMARRSODLHGCACRALVDELENIAR VDPKRTIOMGSEFINPDQSVEVPAARSALHTLELEVCDDMKEGEIDPSTHR KNYVAVSRNCESSLIDGIRIDISIDSLTGFACESIVEEYEDLEIEFFSREADNYK DKLSKRTDLCDDHALHSHDEL"

BASE COUNT 227 a 182 c 228 g 161 t  
 ORIGIN  
 Query Match 73.3%; Score 590.4; DB 10; Length 798;  
 Best Local Similarity 87.0%; Pred. No. 1.1e-141;  
 Matches 672; Conservative 1; Mismatches 92; Indels 7; Gaps 2;

18 CAAAGTGAAGTCCAGCGGCTCTCCAGCGGCTCTCCAGCGGCTTGGGCGACGAA 77  
 29 CAAAGTGAAGTCCAGCGGCTCTCCAGCGGCTCTCCAGCGGCTTGGGCGACGAA 88  
 78 GGTGAGCAACCCCGTTACCTAAATGAAGAGTGGGCTGGGCTGGGCTGGGCTGG 137  
 89 GGTGAGCAACCCCGTTACCTAAATGAAGAGTGGGCTGGGCTGGGCTGGGCTGG 148  
 138 GGGGCGCTGGGAGAACCGGCTGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAG 197  
 149 GGGGCGCTGGGAGAACCGGCTGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAG 208  
 198 GGGGCTGGTGGATGAATGAATGGGAAATGGCCAGTGGAGGAGGAGGAGGAG 257  
 209 GGGGCTGGTGGATGAATGAATGGGAAATGGCCAGTGGAGGAGGAGGAGGAG 268  
 258 AGATGGATCTTCCGAGATCAATGAGATGGAGGAGGAGGAGGAGGAGGAGGAG 317

Db	269	AGATGGATTCCTCCGAAATCAATCCAGATGGCAGCCAGTACGTTGTGGAGTACTTANG	328
OY	318	CCCCGTGAGAGGCCCAACCTCAACAGACGTCTGCGAGAGATATGTACACCGGATGAAGAGCT	377
Db	329	CCCCGTGAGAGGCCCAACCTCAACAGAGTGTGTGAGAGAGGTGTGTACCCGAATGAAGAGCT	388
OY	378	ATGGGGAACAGATTGATCTCTTCACCACATGCGAAGACTACTAGCTGTAAGTGGGCCGGA	437
Db	389	ACGGGGAACAGATTGATCTCTTCACCACATGCGAAGACTACTAGCTGTAAGTGGGCCGGA	448
OY	438	ATGGAGAAATCCAGTGAATCTGGACCTTCAACAGGATATCCGAATGCATCTCAGATATTACCGCA	497
Db	449	ATGGAGAAATCCAGTGAATCTGGACCTTCAACAGGATATCCGAATGCATCTCAGATATTACCGCA	508
OY	488	CCCTCAAGTTTGCCTGTGAGAGCATTTGTGGAGGAATTAGAGAGTGAACATCTGAATTTCT	557
Db	509	CCCTCAAGTTTGCCTGTGAGAGCATTTGTGGAGGAATTAGAGAGTGAACATCTGAATTTCT	568
OY	558	TTTCCCGAGAGGCTTACAATGTTAAAGCAAACTTTGCAGTAAAGCAACAGATCTTTGTG	617
Db	569	TCTCCAGAGAGGCTTACAACGTTAAAGCAAACTTTGCAGTAAAGCGAGCATGTCATGTG	628
OY	618	ACCATGACCCCTGCATATATTCGCATGTATGTAGCTTGAACACATGAGACAGCCAC-----AC	672
Db	629	ACCATGACCCCTGCATATATTCGCATGTATGTAGCTTGAACACATGAGACAGCCCTTACACC	688
OY	673	TGGCTTGATGATCACACCCCAAGAGAGGGGAAAAATGCTGCAATGCTTTTATATTTATATGT	732
Db	689	AAACCTGATGTCACACCCCAAGAGAGGGGAAAGATGACCATGTCCTTT--TATATTACGT	746
OY	733	TTTTTGTGAATTAATCTGAAAAAATATGAAACCAAAAGTAAAAAATAAAAA 784	
Db	747	TTTTTGTGAATTAATCTGAAAAAATCTCTTGAACCCGAAAAAATAAAAA 798	

LOCUS	158198 bp	DNA	linear	HTG 17-JUL-2001
DEFINITION	Homo sapiens chromosome 12 clone RP11-764LL4, WORKING DRAFT			
ACCESSION	AC012013			
VERSION	AC012013.18 GI:14717292			
KEYWORDS	HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEEYIN.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Muzny,D.M., Adams,C., Ado-Oduola,B., Ali-osman,F.R., Allen,C., Albrook,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Butch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elahi,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Gasta,Z., Gill,R., Gorrell,J.H., Guevarez,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Hawla,K., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Hulys,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivel,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martinale,A., Martinez,E., Massey,E., McWhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzner,M., Milner,G., Milner,Z., Mitchell,T., Mohabbat,K.,			

Moragan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
 Nguyen, N., Nguyen, N., Nguyen, N., Nickerson, E., Nwokweto, S.,  
 Ogih, M., Okunolu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B.,  
 Peery, J., Peters, L., Peters, L., Pickens, R., Primeus, E., Pu, L.,  
 Qulles, M., Ren, Y., Rives, M., Rojas, A., Rojiboken, I., Rolfe, M.,  
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohart, N.,  
 Slason, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Wallington, S., Williams, G., Williamson, A., Wleczky, R., Woodson, S.,  
 Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 158198)  
 Worley, K.C.  
 Direct Submission  
 Submitted (19-OCT-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 13, 2001 this sequence version replaced gi:14547727.  
 -----  
 Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 -----  
 Project Information  
 Center project name: HMKZ  
 Center clone name: RP11-764L14  
 -----  
 Summary Statistics  
 Sequencing vector: M13; 108821  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 161278 bases at least 940  
 Consensus quality: 170806 bases at least 920  
 Consensus quality: 174843 bases at least Q20  
 Estimated insert size: 166803; sum-of-contigs estimation  
 Estimated insert size: 163288; agarose-fp estimation  
 Quality coverage: 5x in Q20 bases; agarose-fp estimation  
 Quality coverage: 5.8x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 108229: contig of 108229 bp in length  
 \* 108329: gap of unknown length  
 \* 108330 108246: contig of 41917 bp in length  
 \* 150247 150346: gap of unknown length  
 \* 150347 153410: contig of 3064 bp in length  
 \* 153411 153510: gap of unknown length  
 \* 153511 155739: contig of 2229 bp in length  
 \* 155740 155839: gap of unknown length  
 \* 155840 158198: contig of 2359 bp in length.  
 Location/Qualifiers  
 1. 158198  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-764L14"  
 -----  
 BASE COUNT 46057 a 30922 c 31747 g 49030 t 442 others  
 ORIGIN  
 Query Match 73.3%; Score 590.4; DB 2; Length 158198;  
 Best Local Similarity 87.0%; Pred. NO. 3.2e-141;

[illegible]

AUTHORS	Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.
TITLE	Direct Submision
JOURNAL	Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA
FEATURES	
SOURCE	Location/Qualifiers 1..1017 /organism="Mus musculus" /db_xref="dbest:AA049839" /db_xref="taxon:10090" /clone="IMAGE:476061" 1..1017 /gene="Zs19"
gene	/gene="Zs19"
CDS	343..891 /gene="Zs19" /codon_start=1 /product="putative secreted protein Zs19"
BASE COUNT	248 a     250 c     315 g     204 t
ORIGIN	
Query Match	72.4%; Score 583.2; DB 10; Length 1017;
Best Local Similarity	88.3%; Pred. No. 8e-140;
Matches	657; Conservative 1; Mismatches 79; Indels 7; Gaps 2;
Dn	18 CAAATGAAAGTCACAGCGGTCTGCCAGCGCCTTGGCCACGGCGCGCCCTGGACCAA 77
Dc	257 CAAATGAGAGTCGCGCTGCTTCCAGAGCCCTGGCCACGGCGCGCCCTGGAGCAGA 316
Dy	78 GGTGAGCAACCCGTTACCCTTAARATAAAGCTTGAGGTTGGCTGGCCCTGCCTTC 137
Dd	317 GGTCGCTGCTGGGAAGTGGCTGGCTCGAAGAGCCAAAGATCTACATGTTGAGCTTGA 436
Dq	138 GGGCCCTGCTGGGAACCGCCTGGGCTCGAGAGAGCCAGATCTCCACTGTGAGCATCA 197
Df	377 GGGCTCTGCTGGGAAGTGGCTGGCTCGAAGAGCCAAAGATCTACATGTTGAGCTTGA 436
Dg	198 GGGCTCTGCTGGATGACTAGAAATGGGAATTGCCAGGTGACCCCAAGAAGACATTC 257
Dh	437 GGGCTCTGCTGGATGAAATTAGAGTGGGAATTGCCCGCTGAGACCCCAAGAAGACATTC 496
Di	258 AGATGAGATCTTTCCGGATCAATCCAGATGGAGCAGCTACGTGGTGGAGGCGCTTATG 317
Dj	497 AGATGAGATCTTTCCGGATCAATCCAGATGGAGCAGCTACGTGGTGGAGGCGCTTATG 556
Dk	318 CCCGCTCAGAGCCACCTCACAGAGACTGCTGGAGAGATATGTGACCGATGAAGAGT 377
Dl	557 CCCGCTCAGAGCCACCTCACAGAGAGTGTGTTGAGAGAGTGTGTGACCGATGAAGAGT 616
Dm	378 ATGGGAAACAGATTGATCTTCCACCCATTCGACAGACTACGATAGTGTAGTGGCGGA 437
Dn	617 ACGGGGAACAGATTGATCTTCCACCCATTCGACAGACTACGATAGTGTAGTGGCGGA 676
Dp	438 ATGGAGATCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 497
Dq	677 ATGGAGATCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 736
Dr	498 CCTTCAGTTGCTGCTGAGAGCATTTGAGAGATAGAGAGATAGATCAATTGAAATCT 557
Ds	737 CCTTCAGTTGCTGCTGAGAGCATTTGAGAGATAGAGAGATAGATCAATTGAAATCT 796
Dt	558 TTTCCGAGAGGCTGACAATGTTAAAGCAAACTTTGAGTAGTAAGCAAGACAGATCTTGTG 617
Du	797 TTTCCGAGAGGCTGACAATGTTAAAGCAAACTTTGAGTAGTAAGCAAGACAGATCTATGTG 856
Dv	618 ACATGCCCCGTGACATATTCGATATGATGATATGAAACACATGAGAGACCCAC-----AC 672
Dw	857 ACCATGCCCCGTGACATATTCGATATGATGATATGAAACACATGAGAGACCCAC-----AC 916

```

OY 673 TGGCTGATGATCACCAGAGGAGGAAATGTGGCAATGCTTTATATATATATGT 732
DB 917 AAGCGTATGGAACACCCAGAGGAGGAGATGCAGATGCTCTT--TATATTACGT 974
OY 733 TTTTACGAATTAACGAAAAA 756
DB 975 TTTTATGCAATGAACTGAAAAA 998

RESULT 8
AF186113 649 bp mRNA linear PRI 13-JAN-2000
LOCUS Homo sapiens putative secreted protein zsig9 (ZSIG9) mRNA, complete cds.
ACCESSION AF186113
VERSION AF186113.1 GI:6014631
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 649)
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.
TITLE Homo sapiens putative secreted protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 649)
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA
FEATURES
source
1..649
/organism="Homo sapiens"
/db_xref="dbEST:RI5306"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q15"
/clone="IMAGE:29675"
1..649
/gene="ZSIG9"
104..358
/gene="ZSIG9"
/codon_start=1
/product="putative secreted protein ZSIG9"
/protein_id="AA01431.1"
/db_xref="GI:6014632"
/translation="MKGMGLALLGLALGTAMARRSODLHGACRALVDELEMEIAQVDPKTIOMGSFRINPDGOSVVEVTVPKRVHSGFG"
BASE COUNT 175 a 150 c 171 g 149 t 4 others
ORIGIN
Query Match 38.3%; Score 308.6; DB 9; Length 649;
Best Local Similarity 98.1%; Pied. No. 6.4e-69;
Matches 312; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CGGCCCCAAGCTGGGCGCAAGTGAAGTGCAGGGCTCTGGGCGCACGGCG 60
DB 1 CGGCCCCAAGCTGGGCGCAAGTGAAGTGCAGGGCTCTGCCAGCGCTGGGCGCACGGCG 60
OY 61 GCGGCCCTGGGACCAAGGTGAGCAACCCGTTACCTTAARATGAAGAAGCTGGGTTG 120
DB 61 GCGGCCCTGGGACCAAGGTGAGCAACCCGTTACCTTAARATGAAGAAGCTGGGTTG 120
OY 121 GCTGGCCCTCTTCTGGGGGCGCTGCTGGGACCCGCTGGGCTGCGAGAGCCAGATCT 180
DB 121 GCTGGCCCTCTTCTGGGGGCGCTGCTGGGACCCGCTGGGCTGCGAGAGCCAGATCT 180
OY 181 CCACTGTGAGCATGCAAGGCTCTGGTGAATGAATGGAATTCGCCAGGTGGA 240
DB 181 CCACTGTGAGCATGCAAGGCTCTGGTGAATGAATGGAATTCGCCAGGTGGA 240

```

```

OY 241 CCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCAGATGCAGCCAGTCACT 300
DB 241 CCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCAGATGCAGCCAGTCACT 300
OY 301 GGTGAGGTGCTCTTATGC 318
DB 301 GGTGAGGTGCTCTTATGC 318

RESULT 9
BC001027 824 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, transmembrane protein 4, clone MGC:1545
DEFINITION IMAGE:3344788, mRNA, complete cds.
ACCESSION BC001027
VERSION BC001027.1 GI:12654402
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 824)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
REMARK COMMENT
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gatherersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M., Benjamini,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Guan,X., Gupta,V., Ho,S.-L., Karlins,E., Legaspi,R., Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Snyder,B., Stantirpop,S., Thomas,P.J., Tjongson,E.E., Touchman,J.W., Tsugeon,C., Vogt,J.L., Walker,M.A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 5 Row: P Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6014631.
FEATURES
source
1..824
/organism="Homo sapiens"
/db_xref="locustid:10330"
/db_xref="taxon:9606"
/clone="MGC:1545 IMAGE:3344788"
/cisue_type="Eye, retinoblastoma"
/clone_lib="NIR_MGC_16"
/lab_host="DH10B-R"
/notes="Vector: pOT87"
250..504
/codon_start=1
/product="transmembrane protein 4"
/protein_id="AAH01027.1"
/db_xref="GI:12654403"
/translation="MKGMGLALLGLALGTAMARRSODLHGACRALVDELEMEIAQVDPKTIOMGSFRINPDGOSVVEVTVPKRVHSGFG"
BASE COUNT 221 a 194 c 233 g 176 t
ORIGIN

```

Query Match 35.8%; Score 288.4; DB 9; Length 824;  
 Best Local Similarity 95.8%; Pred. No. 1.1e-63;  
 Matches 295; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

11 CTTGGGCCCCAAATGAAAGCCGAGCCGCTGCGAGCGCTGGGCGCCGCGGCGCCCTG 70  
 |||||||  
 157 CTTGGGCCCCAAATGAAAGCCGAGCCGCTGCGAGCGCTGGGCGCCGCGGCGCCCTG 216  
 |||||||

71 GACCAAGGTGGAGCAACCCCGCTACCCCTTAARATGAAAGGCTGGGCTGGCTGGCCCTG 130  
 |||||||  
 217 GAGCAGAGGTGGAGCAACCCCATTAAGCTTAAGATGAAAGGCTGGGCTGGCTGGCCCTG 276  
 |||||||

131 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGACCCAGATGTCACATGTGGA 190  
 |||||||  
 277 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGAGCCAGATGTCACATGTGGA 336  
 |||||||

191 GCATGAGGGCTCTGTGTGATGACTAGAAATGGAAATGGCCAGGTGGAGCCCAAGAG 250  
 |||||||  
 337 GCATGAGGGCTCTGTGTGATGACTAGAAATGGAAATGGCCAGGTGGAGCCCAAGAG 396  
 |||||||

251 ACCATTACAGATGGATCTTCCGATCATCATGAGCGACGACATGCTGGTGGAGTGTG 310  
 |||||||  
 397 ACCATTACAGATGGATCTTCCGATCATCATGAGCGACGACATGCTGGTGGAGTGTG 456  
 |||||||

311 CCTTATGC 318  
 |||  
 457 ACTGTTC 464  
 |||

RESULT 10  
 AX440375 564 bp DNA linear PAT 28-JUN-2002  
 LOCUS Sequence 228 from Patent WO0190154.  
 DEFINITION AX440375  
 ACCESSION AX440375  
 VERSION AX440375.1 GI:21651185  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H.,  
 Luder, M.J., Algate, P.A., Pfling, S.P., Mannion, J., Benson, D.R. and  
 Carter, D.  
 TITLE Compositions and methods for the therapy and diagnosis of ovarian  
 cancer  
 JOURNAL Patent: WO 0190154-A 228 29-NOV-2001;  
 CORIAX CORPORATION (US)  
 FEATURES  
 source location/Qualifiers  
 1. 564  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 147 a 142 c 152 g 118 t 5 others  
 ORIGIN

Query Match 29.5%; Score 237.8; DB 6; Length 564;  
 Best Local Similarity 92.4%; Pred. No. 1.2e-50;  
 Matches 281; Conservative 1; Mismatches 13; Indels 9; Gaps 3;

11 CTTGGGCCCCAAATGAAAGCCGAGCCGCTGCGAGCGCTGGGCGCCGCGGCGCCCTG 70  
 |||||||  
 269 CTTGGGCCCCAAATGAAAGCCGAGCCGCTGCGAGCGCTGGGCGCCGCGGCGCCCTG 328  
 |||||||

71 GACCAAGGTGGAGCAACCCCGCTACCCCTTAARATGAAAGGCTGGGCTGGCTGGCCCTG 130  
 |||||||  
 329 GAGCAGAGGTGGAGCAACCCCATTAAGCTTAAGATGAAAGGCTGGGCTGGCTGGCCCTG 388  
 |||||||

131 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGACCCAGATGTCACATGTGGA 190  
 |||||||  
 389 CTTTGGGGGGCCCTGCTGGGAACCCGCTGGCTCGAGAGACCCAGATGTCACATGTGGA 447  
 |||||||

191 GCATGAGGGCTCTGTGTGATGACTAGAAATGGAAATGGCCAGGTGGAGCCCAAGAG 250  
 |||||||

DB 448 -----AGGCGCTGTGTGATGACTAGAAATGGAAATGG-CCANGTGAGCCCAAGAG 500  
 QY 251 ACCATTACAGATGGATCTTCCGATC-AATCCAGATGAGCAGATGCTGGTGGAGGT 309  
 |||||||  
 DB 501 ACCATTACAGATGGATCTTCCGATC-AATCCAGATGAGCAGATGCTGGTGGAGGT 560  
 |||||||

QY 310 GCCT 313  
 |||  
 DB 561 GCCT 564

RESULT 11  
 AC025574 155023 bp DNA linear HTG 09-MAY-2002  
 LOCUS Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT  
 DEFINITION AC025574  
 ACCESSION AC025574  
 VERSION AC025574.13 GI:20428723  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 155023)  
 Muzny, D.M., Adams, C., Adlo-Oduola, B., Ali-Osman, F.R., Allen, C.,  
 Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Ayale, M., Banks, T.,  
 Barbara, J., Benton, J., Blimie, K., Blankenburg, K., Bonin, D.,  
 Bouck, J., Boyle, S., Brileva, M., Brown, E., Brown, N., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,  
 Homsl, F., Howard, S., Huber, J., Hulys, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, M., Louised, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mathew, E., McLeod, M.P., Meador, M., Mei, G., Metker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokkenko, S., Ogul, M., Okunolu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Qules, M., Ren, Y.,  
 Rivers, M., Rojas, A., Rojuben, I., Rolfe, M., Ruiz, S., Severy, G.,  
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,  
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tameis, A., Tameis, K., Tang, H.,  
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstock, G. and Gibbs, R.

Direct Submission  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 155023)  
 AUTHORS Worley, K.C.  
 REFERENCE Direct Submission  
 Submitted (11-MAR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 155023)







JOURNAL Submitted (01-JUL-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 REFERENCE 3 (bases 1 to 182350)  
 AUTHORS Morley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 COMMENT On Aug 23, 2002 this sequence version replaced gi:22296903.

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Drafting Center Code: BCM  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: HBWI  
 Center clone name: RP11-977G19  
 Summary Statistics  
 Sequencing vector: Plasmid  
 Sequencing vector: M13  
 Chemistry: Dye-terminator Big Dye 3.1  
 Assembly program: Phrap, version 0.990329  
 Consensus quality: 179279 bases at least Q40  
 Consensus quality: 180581 bases at least Q30  
 Consensus quality: 181073 bases at least Q20  
 Estimated insert size: 174258; sum-of-coverage estimation  
 Quality coverage: 8x in Q20 bases; sum-of-coverage estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 2893: contig of 2893 bp in length  
 \* 2894 2893: gap of unknown length  
 \* 2994 6686: contig of 3693 bp in length  
 \* 6687 6786: gap of unknown length  
 \* 17251 17251: contig of 10465 bp in length  
 \* 17252 17351: gap of unknown length  
 \* 17352 85894: contig of 68543 bp in length  
 \* 85895 85894: gap of unknown length  
 \* 182350 182350: contig of 96356 bp in length.

FEATURES  
 source  
 1.182350  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-977G19"

BASE COUNT 46453 a 45416 c 43855 g 46174 t 452 others  
 ORIGIN

Query Match 25.6% Score 206; DB 2; Length 182350;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-42;  
 Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 306 AGGTGCTTATGCGCCCTCAGAGCCACCTCAGAGGCTGTGAGAGATGTGACC 365  
 |||||||  
 DB 172029 AGGTGCTTATGCGCCCTCAGAGCCACCTCAGAGGCTGTGAGAGATGTGACC 171970  
 |||||||  
 OY 366 GGATGAGAGAGTATGGGGAACAGATTGATCTTCCACCCATCGCAAGAACTACGTCG 425  
 |||||||  
 DB 171969 GGATGAGAGAGTATGGGGAACAGATTGATCTTCCACCCATCGCAAGAACTACGTCG 171910  
 |||||||  
 OY 426 TACTGGCCGGGATGAGATGCAATCGAATCGAATCGAATCGAATCGAATCGAATCGA 485  
 |||||||

DB 171909 TACTGGCCGGGATGAGATGCAATCGAATCGAATCGAATCGAATCGAATCGA 171850  
 OY 486 ATATTAGCGGACACCTCAAGTTTGGC 511  
 |||||||  
 DB 171849 ATATTAGCGGACACCTCAAGTTTGGC 171824

RESULT 13  
 AC023500 184762 bp DNA linear HTG 24-AUG-2002  
 LOCUS Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
 DEFINITION SEQUENCE, 9 unordered pieces.  
 AC023500  
 VERSION AC023500.32 GI:22095070  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-ozman, F.R., Allen, C., Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T., Barbata, J., Benton, J., Blum, K., Blankenburg, K., Bonin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Cartron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, M.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Eamhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homi, F., Howard, S., Huber, J., Hui, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, R., Louised, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mel, G., Metker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Nguyen, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunolu, G., Oregunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Ruiz, S., Savary, G., Rivers, M., Rojas, A., Rojudo, I., Rolfe, M., Rull, M., Ren, Y., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tameris, A., Tameris, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watling, S., Williams, G., Williamson, A., Wleczek, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 184762)  
 AUTHORS Morley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 REFERENCE 3 (bases 1 to 184762)  
 AUTHORS Morley, K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One



OY 426 TAGTGGCCGCGATGAGATCCAGTGAACCTGAGCCTACAGGACATCCGATCGACTGAC 485  
 DB 75907 TCGTGGCCGCGATGAGATCCAGTGAACCTGAGCCTACAGGACATCCGATCGACTGAC 75966  
 OY 486 ATATTAGCGCGACCCCTCAGTTTGGC 511  
 DB 75967 ATATCAGCGCGACCCCTCAGTTTGGC 75992

RESULT 15  
 AC109891 101666 bp DNA linear HTG 13-JUL-2002  
 LOCUS Rattus norvegicus clone CH230-31986, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 47 unordered pieces.  
 AC109891  
 VERSION AC109891.3 GI:21738196  
 KEYWORDS HTG; HTGS PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Ratus.  
 1 (bases 1 to 101666)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayala,M., Banks,T.,  
 Barbarta,J., Benton,J., Blamey,K., Blankenburg,K., Bonni,D.,  
 Bouck,J., Bowles,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Delaney,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Hawlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
 Homel,F., Howard,S., Huber,J., Huliy,K., Hummel,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratochvil,A., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Liew,C., Liu,J., Liu,W., Louissege,H.,  
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzer,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okuwon,G.,  
 Ogunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Qulles,M., Ren,Y.,  
 Rivers,M., Rojas,A., Rojubenok,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Syatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Umanil,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,O.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williams,A., Wleczek,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G., and Gldbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 101666)  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 3 (bases 1 to 101666)

AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Worley,K.C.  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced g1.18860222.  
 ----- Genome Center -----  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information -----  
 Center project name: GSKY  
 Center clone name: CH230-31986  
 ----- Summary Statistics -----  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 60229 bases at least Q40  
 Consensus quality: 63959 bases at least Q30  
 Consensus quality: 66730 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a "working draft" sequence. It currently  
 \* consists of 47 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1  
 \* 1423: contig of 1423 bp in length  
 \* 1424 1523: gap of unknown length  
 \* 1524 2972: contig of 1449 bp in length  
 \* 2973 3072: gap of unknown length  
 \* 3073 4717: contig of 1645 bp in length  
 \* 4718 4817: gap of unknown length  
 \* 4818 6633: contig of 1815 bp in length  
 \* 6633 6732: gap of unknown length  
 \* 6733 8040: contig of 1308 bp in length  
 \* 8041 8140: gap of unknown length  
 \* 8141 10043: contig of 1903 bp in length  
 \* 10044 10143: gap of unknown length  
 \* 10144 11337: contig of 1094 bp in length  
 \* 11338 11337: gap of unknown length  
 \* 11338 12362: contig of 1025 bp in length  
 \* 12363 12462: gap of unknown length  
 \* 12463 14100: contig of 1638 bp in length  
 \* 14101 14200: gap of unknown length  
 \* 14201 15840: contig of 1640 bp in length  
 \* 15841 15940: gap of unknown length  
 \* 15941 17053: contig of 1113 bp in length  
 \* 17054 17153: gap of unknown length  
 \* 17154 18836: contig of 1683 bp in length  
 \* 18837 18936: gap of unknown length  
 \* 18937 20137: contig of 1201 bp in length  
 \* 20138 20237: gap of unknown length  
 \* 20238 21737: gap of unknown length  
 \* 21737 21836: gap of unknown length  
 \* 21837 23519: contig of 1783 bp in length  
 \* 23520 23719: gap of unknown length  
 \* 23720 25907: contig of 2188 bp in length  
 \* 25908 27406: contig of 1399 bp in length  
 \* 27407 27506: gap of unknown length  
 \* 27507 28837: contig of 1331 bp in length  
 \* 28838 28937: gap of unknown length  
 \* 28938 30329: contig of 1292 bp in length  
 \* 30330 32012: gap of unknown length  
 \* 32013 32112: contig of 1683 bp in length  
 \* 32113 33198: gap of unknown length  
 \* 33198: contig of 1086 bp in length

```

* 33199 33298: gap of unknown length
* 33299 35519: contig of 2221 bp in length
* 33520 35619: gap of unknown length
* 35620 37677: contig of 2058 bp in length
* 37678 37777: gap of unknown length
* 37778 38936: contig of 1159 bp in length
* 38937 39036: gap of unknown length
* 39037 40273: contig of 1237 bp in length
* 40274 40373: gap of unknown length
* 40374 41906: contig of 1533 bp in length
* 41907 42006: gap of unknown length
* 42007 44506: contig of 2500 bp in length
* 44507 44606: gap of unknown length
* 44607 46803: contig of 2197 bp in length
* 46804 46903: gap of unknown length
* 46904 49226: contig of 2323 bp in length
* 49227 49326: gap of unknown length
* 49327 51433: contig of 2107 bp in length
* 51434 51533: gap of unknown length
* 51534 53763: contig of 2230 bp in length
* 53764 53863: gap of unknown length
* 53864 56613: contig of 2750 bp in length
* 56614 56713: gap of unknown length
* 56714 58773: contig of 2060 bp in length
* 58774 58873: gap of unknown length
* 58874 61527: contig of 2654 bp in length
* 61528 61627: gap of unknown length
* 61628 63192: contig of 1565 bp in length
* 63193 63292: gap of unknown length
* 63293 65362: contig of 2070 bp in length
* 65363 65462: gap of unknown length
* 65463 67438: contig of 1976 bp in length
* 67439 67538: gap of unknown length
* 67539 70505: contig of 2967 bp in length
* 70506 70605: gap of unknown length
* 70606 74524: contig of 3919 bp in length
* 74525 74624: gap of unknown length
* 74625 76737: contig of 2113 bp in length
* 76738 76837: gap of unknown length
* 76838 80133: contig of 3296 bp in length
* 80134 80233: gap of unknown length
* 80234 83122: contig of 2889 bp in length
* 83123 83222: gap of unknown length
* 83223 86731: contig of 3509 bp in length
* 86732 86831: gap of unknown length
* 86832 89929: contig of 3098 bp in length
* 89930 90029: gap of unknown length
* 90030 93107: contig of 3078 bp in length
* 93108 93207: gap of unknown length
* 93208 97419: contig of 4212 bp in length
* 97420 97519: gap of unknown length
* 97520 101666: contig of 4147 bp in length.
FEATURES
source
1.101666
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-319E6"
BASE COUNT 24650 a 22934 c 23569 g 23059 t 7454 others
ORIGIN
Query Match 21.68; Score 174; DB 2; Length 101666;
Best Local Similarity 90.38; Pred. No. 1.1e-33;
Matches 186; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
OY 306 AGGTGCTTATGCCGCTCAGAGCCGACCTCAGAGAGTGTGAGATATGAGC 365
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 97570 AGGTTCTTATGCCGCTCAGAGCCGACCTCAGAGAGTGTGAGATATGAGC 97629
OY 366 GGATGAAGGATGAGGAGACAGATTGATCTTCACCACTGCGAGAACTACGATG 425
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 97630 GAATGAAGGATGAGGAGACAGATTGATCTTCACCACTGCGAGAACTACGATG 97689
OY 426 TAGTGGCCGGAGATGAGATCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 485

```

```

Db 97690 TTGTGCGCCGCGAGAGGAGATCCAGTGAAGTGAAGTGAAGTGAAGTGAAG 97749
OY 486 ATATTAGCGCACCCCTCAGATTGCG 511
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 97750 ATATCACTGCGCACCCCTCAGATTGCG 97775

```

Search completed: March 30, 2003, 20:19:59  
 Job time : 2658.55 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 17:35:17 ; Search time 191.291 Seconds  
(without alignments)  
9488.757 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806  
Sequence: 1 cggcccaagctggtggccaa.....aagagagagagagagacta 806

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

N\_Geneseq.101002:\*  
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	805.6	100.0	806	21	AAZ08293
2	804.6	99.8	806	20	AAZ06969
3	750.4	93.1	814	21	AAZ38327
4	750.4	93.1	814	24	ABK52765
5	750.4	93.1	814	24	ABK41995
6	750.4	93.1	814	24	ABK09772
7	749.4	93.0	1180	20	AAZ22112
8	749.4	93.0	1210	22	AAZ21330
9	702.4	87.1	714	20	AAZ97884

10	688.4	85.4	832	20	AAZ97837
11	646.4	80.2	657	24	ABK36007
12	599	74.3	1069	20	AAZ06970
13	599	74.3	1069	21	AAZ08294
14	546	67.7	546	21	AAZ38326
15	439.2	54.5	592	21	AAZ35325
16	334.8	41.5	558	22	AAZ18023
17	318.6	39.5	595	22	AAZ07882
18	308.6	38.3	649	20	AAZ06968
19	306.8	38.1	649	21	AAZ08284
20	288	35.7	718	20	AAZ10655
21	286	35.5	792	20	AAZ97836
22	284.2	35.3	1085	20	AAZ22130
23	279.2	34.6	649	20	AAZ97883
24	272.4	33.8	484	22	AAZ25628
25	271.4	33.7	1160	22	AAZ26689
26	267.6	33.2	415	20	AAZ06971
27	267	33.1	415	21	AAZ08285
28	237.8	29.5	564	24	ABK09691
29	173.8	21.6	352	22	ABK09488
30	163.8	20.3	215	21	AAZ24962
31	151.8	18.8	265	20	AAZ39549
32	147	18.2	147	24	ABN94837
33	147	18.2	147	24	ABL65469
34	130.8	16.2	495	22	AAZ180994
35	126	15.6	126	24	ABN94836
36	126	15.6	126	24	ABL65468
37	121	15.0	2532	22	ABA07341
38	121	15.0	2532	22	AAZ32768
39	121	15.0	2724	22	ABK07340
40	121	15.0	2724	22	AAZ32767
41	60	7.4	60	24	ABN39205
42	53.8	6.7	65	24	ABN55235
43	53.4	6.6	816	23	ABE21471
44	51.4	6.4	3110	23	ABL21470
45	50.4	6.3	483	23	ABV59155

#### ALIGNMENTS

RESULT 1	AAZ08293	AAZ08293 standard; DNA; 806 BP.
ID	AAZ08293	
XX	AAZ08293	
AC	AAZ08293	
XX		
DT	07-FEB-2000	(first entry)
XX		
DE	Human Zs199 gene encoding secretory protein variant-4.	
XX		
KW	Secretory protein-9; Human Zs199; chromosome 12q15 region; variant;	
KW	overexpression; antagonist; antibody; antisense nucleotide; tumour;	
KW	treatment; receptor; radio-label; fusion; polypeptide toxin; technique;	
KW	down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;	
KW	detection; stomach; lymphoma; alternative splicing; allelic variation;	
KW	silent mutation; ds.	
XX		
OS	Homo sapiens.	
XX		
XX	Synthetic.	
FT	Key	Location/Qualifiers
FT	CDS	104..652
FT		/tag= a
FT		/product= "Zs199 secretory protein variant-4"
FT		/note= "Overexpressed in tumours"
FT		/tag= b
FT	sig_peptide	104..163
FT		/tag= c
FT	mat_peptide	164..649
FT		/tag= c
FT		/label= Mature_Zs199_Protein_variant-4
PN	WO960405-A1.	

XX	25-NOV-1999.
PD	
XX	
PE	19-MAY-1999; 99WO-US11107.
XX	
PR	19-MAY-1998; 98US-0081183.
XX	
PA	(Zymo ) ZYMOGENETICS INC.
PI	Moore EE, Taft DW;
DR	WPI; 2000-039447/03.
P-PSDB:	AAV15135.
XX	
PT	Detecting tumors using antibodies, antagonists and antisense
XX	nucleotides to secretory protein-9 (Zs1g9) -
PS	
XX	Disclosure; Page 33-35; 45pp; English.
CC	
CC	The present DNA sequence is a gene encoding the variant of the secretory
CC	protein-9, Zs1g9 that arises due to alternative splicing, allelic
CC	variation or silent mutations that result in amino acid changes. This
CC	sequence is mapped to the human chromosome 12q15 region. It is
CC	overexpressed in tumours. Antagonists, antibodies and antisense
CC	nucleotides to Zs1g9 are useful for detecting and treating tumors. The
CC	antagonist may be an antibody or receptor to Zs1g9 and it may be radio-
CC	labelled or fused to a polypeptide toxin. It can be used for down
CC	regulating the overexpression of Zs1g9. The gene sequence can be used as
CC	nucleic acid probes to detect RNA encoding Zs1g9. The Zs1g9 sequence
CC	facilitates improved diagnostic and therapeutic techniques for detecting
CC	and treating cancers, especially of the brain, liver, stomach, lymphoma,
CC	etc., at an early stage.
XX	
SQ	Sequence 806 BP; 229 A; 183 C; 223 G; 160 T; 1 other;
	Query Match 100.0%; Score 805.6; DB 21; Length 806;
	Best Local Similarity 100.0%; Pred. No. 1,2e-193;
	Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CGGCCCAAGCTGGGGCCAAAGTGAAGTCACGCGGTGCCAGCGCTTGCGGCACGCGC 60
Db	1 CGGCCCAAGCTGGGGCCAAAGTGAAGTCACGCGGTGCCAGCGCTTGCGGCACGCGC 60
OY	61 GCGGCCCTGGGACCAAAGTGAAGCACACCCTTACCTTAARATGAAGAAGCTGGGGTTG 120
Db	61 GCGGCCCTGGGACCAAAGTGAAGCACACCCTTACCTTAARATGAAGAAGCTGGGGTTG 120
OY	121 GCTGCCCCCTGTCTTGGGGGCCCTCTGCGGAAACGCGTGGGGCTGGAGAGCCAGGATCT 180
Db	121 GCTGCCCCCTGTCTTGGGGGCCCTCTCTGGGAACCCCTTGGGCTGGAGAGCCAGGATCT 180
OY	181 CCACGTGTGAGCATGCAGAGGCTCTGTGGATGAACTAGAAATGGAAAATTGCCACAGTGA 240
Db	181 CCACGTGTGAGCATGCAGAGGCTCTGTGGATGAACTAGAAATGGAAAATTGCCACAGTGA 240
OY	241 CCCCAAGAAGACCAATTCAGATGGGATCTTTCGCGATCAATCCAAGATGGCAGCAGTCACT 300
Db	241 CCCCAAGAAGACCAATTCAGATGGGATCTTTCGCGATCAATCCAAGATGGCAGCAGTCACT 300
OY	301 GGATGAGTGGCTTATGGCCGCTCAGAGGCCCACTCAAGAGCTGTGGAGAGATATG 360
Db	301 GGATGAGTGGCTTATGGCCGCTCAGAGGCCCACTCAAGAGCTGTGGAGAGATATG 360
OY	361 TGACCGGATGAAGAGATATGGGGACAACATATGATCTTCCACCCATTCGCAGAACTACGT 420
Db	361 TGACCGGATGAAGAGATATGGGGACAACATATGATCTTCCACCCATTCGCAGAACTACGT 420
OY	421 ACCTGTACTGGGCCGGAATGAGAACATCAAGTGAACCTACAAAGCATCCGAATGA 480
Db	421 ACCTGTACTGGGCCGGAATGAGAACATCAAGTGAACCTACAAAGCATCCGAATGA 480
OY	481 CTCAGATATTAGCGCACCTCAAGTTTGCTGTGAGAGCATTTGTGGAGGAATACGAGA 540
Db	481 CTCAGATATTAGCGCACCTCAAGTTTGCTGTGAGAGCATTTGTGGAGGAATACGAGA 540

Db	481	CTCAGATATTTAGCGGCACACCCCTCAAGTTTGGGTGTGAAGACATGTGTGAGGAATACGAGGA	540
Oy	541	TGAACTCATTTGAATTTCTTTTCCCGAGAGGCTGACATGTTAAAGACAACTTTGCAGTAA	600
Db	541	TGAATCTTCTTAAATTCCTTTTCCCGAGAGGCTGACAAATGTTAAAGACAACTTTGCAGTAA	600
Oy	601	GGAAACAGATCTTTGTGTGACATGCGCCCGACACATATGCGATGAGCTATGAACCACTGG	660
Db	601	GGAAACAGATCTTTGTGTGACATGCGCCCTGACACATATGCGATGAGCTATGAACCACTGG	660
Oy	661	AGCAGCCACACACTGCTTGATGATGATCAACCCCGAGAGGGGAAATGTGTGCAATGCTTT	720
Db	661	AGCAGCCACACACTGCTTGATGATGATCAACCCCGAGAGGGGAAATGTGTGCAATGCTTT	720
Oy	721	TATATATATGTTTTTTTACTGTAATTTAACTGAAAAATATGAAACCAAAAGCTAAAAAAA	780
Db	721	TATATATATGTTTTTTTACTGTAATTTAACTGAAAAATATGAAACCAAAAGCTAAAAAAA	780
Oy	781	AAAAAAAAAGAGAGAGAGAGAGAACTA	806
Db	781	AAAAAAAAAGAGAGAGAGAGAGAACTA	806
RESULT 2			
ID	AAx06969	standard; cDNA; 806 BP.	
XX	AAx06969;		
AC	AAx06969;		
DT	10-MAY-1999	(first entry)	
XX			
DE	Human secretory peptide-9 (Zs199) variant cDNA.		
XX			
.KW.	Secretory peptide-9; Zs199; human; tumour marker; cancer; therapy;		
KW	diagnosis; growth enhancer; variant; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	104..652	
FT		/*tag= a	
FT	sig_peptide	104..163	
FT		/*tag= b	
FT	mat_peptide	164..649	
FT		/*tag= c	
XX			
PN	W09901554-A1.		
XX			
PD	14-JAN-1999.		
XX			
PE	02-JUL-1998;	98WO-US13859.	
XX			
PR	17-JUN-1998;	98US-0099005.	
PR	03-JUL-1997;	97US-0051704.	
PR	03-JUL-1997;	97US-0888088.	
PR	19-MAY-1998;	98US-0081338.	
PR	19-MAY-1998;	98US-0085983.	
PR	17-JUN-1998;	98US-0089899.	
XX			
PA	(ZYMO ) ZYMOGENETICS INC.		
XX			
PI	Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;		
XX			
DR	WPI: 1999-106055/09.		
DR	P-PSDB; AAM88474.		
XX			
PT	New mammalian secretory peptide-9 (Zs199) - used as a growth		
PT	enhancer for placenta, liver and heart, and as an indicator of		
XX	cancer		
PS	Claim 2; Page 73-74; 85pp: English.		
XX			
CC	This cDNA clone encodes human secretory peptide-9, or Zs199,		



CC variant (see AAW88474). Zs19 (see also AAW88469) is overexpressed in  
 CC human brain, liver, lung, oesophageal, stomach, colon, rectal,  
 CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an  
 CC indicator for cancer. Zs19 cDNA was discovered in a placenta  
 CC clone from a full-term pregnancy cDNA library which contained an  
 CC expressed sequence tag (see AAW06971). The invention provides  
 CC polynucleotides (see AAW06968-70) encoding Zs19 polypeptides (see  
 CC AAW88469-77) including mature polypeptides, other processed forms,  
 CC variants and mouse orthologues. The Zs19 gene, or probes derived  
 CC from it, can be used to determine if Zs19 is present on chromosome  
 CC 12, and if a mutation has occurred. Antibodies raised against  
 CC Zs19 can be used as diagnostic agents to determine the presence of  
 CC Zs19, and thus the presence of cancer. They can also be labelled  
 CC with radioisotopes or fused with toxins and used to treat tumours  
 CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.

SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;

Query Match 99.8%; Score 804.6; DB 20; Length 806;  
 Best Local Similarity 99.9%; Pred. No. 2,1e-193;

Matches 805; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGGCCCAAGCTGGGGCCAAAGTCAAGCGCTGCCAGCGCTTGGGCCACGCG 60  
 DB 1 CGGCCCAAGCTGGGGCCAAAGTCAAGCGCTGCCAGCGCTTGGGCCACGCG 60  
 OY 61 GGGGCGCTGGGACCAAGGTGGAGCAACCCCTTACCTTAATATTAAGCTGGGCTTG 120  
 DB 61 GGGGCGCTGGGACCAAGGTGGAGCAACCCCTTACCTTAATATTAAGCTGGGCTTG 120  
 OY 121 GCTGGCGCTGCTTCTGGGGGCGCTGCTGGGAAACCGCTGGGCGGAGAGCAGATCT 180  
 DB 121 GCTGGCGCTGCTTCTGGGGGCGCTGCTGGGAAACCGCTGGGCGGAGAGATCT 180  
 OY 181 CCAGTGTGAGCATGCAAGGCTCTGTGTGATGAATGAATGGGAATTTGCCAGGTGGA 240  
 DB 181 CCAGTGTGAGCATGCAAGGCTCTGTGTGATGAATGAATGGGAATTTGCCAGGTGGA 240  
 OY 241 CCCCAAGAAAGACCATTCATGAGATGCGGATCTTCCGATCAATCCAGATGCCACGATCACT 300  
 DB 241 CCCCAAGAAAGACCATTCATGAGATGCGGATCTTCCGATCAATCCAGATGCCACGATCACT 300  
 OY 301 GGTGAGGTGCTTATGCCGCTCAGAGGCCCACTCAGAGAGCTCTGGAGAGATATG 360  
 DB 301 GGTGAGGTGCTTATGCCGCTCAGAGGCCCACTCAGAGAGCTCTGGAGAGATATG 360  
 OY 361 TGAACGGATGAAGAGATATGGGAGACAGATTGATCTTCCACCCATCGCAGAACTACGT 420  
 DB 361 TGAACGGATGAAGAGATATGGGAGACAGATTGATCTTCCACCCATCGCAGAACTACGT 420  
 OY 421 AGGTGTAGTGGGCGCGGAATGAGAAATCCAGTGAATCGGACCTTAACAGGATCCGATTCGA 480  
 DB 421 AGGTGTAGTGGGCGCGGAATGAGAAATCCAGTGAATCGGACCTTAACAGGATCCGATTCGA 480  
 OY 481 CTCAGATATTAGGGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGGAGGAATACGAGGA 540  
 DB 481 CTCAGATATTAGGGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGGAGGAATACGAGGA 540  
 OY 541 TGAACATGATGAATCTTTTCCCGAGAGGCTGCAATGTTAAAGCAAACTTTGCACTAA 600  
 DB 541 TGAACATGATGAATCTTTTCCCGAGAGGCTGCAATGTTAAAGCAAACTTTGCACTAA 600  
 OY 601 GCGAAGAGATCTTTGTGACATGCGCTGCACATATGCGATGATGACCTTGAACCACTGG 660  
 DB 601 GCGAAGAGATCTTTGTGACATGCGCTGCACATATGCGATGATGACCTTGAACCACTGG 660  
 OY 661 AGCAGCCCACTGCTGCTTATGATGATACCCCGAGGAGGGAATATGTTGGCAATGCTTT 720  
 DB 661 AGCAGCCCACTGCTGCTTATGATGATACCCCGAGGAGGGAATATGTTGGCAATGCTTT 720

OY 721 TATATATATGTTTTTACTGAATAATCAATGAAAAATATGAAACCAAAAGTAAAAA 780  
 DB 721 TATATATATGTTTTTACTGAATAATCAATGAAAAATATGAAACCAAAAGTAAAAA 780  
 OY 781 AAAAAAAGAGAGAGAGAGAGAACTA 806  
 DB 781 AAAAAAAGAGAGAGAGAGAGAACTA 806

# RESULT 3

AAZ38327  
 ID AAZ38327 standard; cDNA; 814 BP.

AAZ38327;

09-FEB-2000 (first entry)

Human transmembrane protein cDNA clone HP10390.

HP10390; transmembrane domain; stomach cancer cell; antibody;

assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;

agonist; antagonist; ligand; therapeutic; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 145..693

FT /tag= a /product= "human transmembrane protein HP10390"

PN WO955862-A2.

PD 04-NOV-1999.

PE 27-APR-1999; 99WO-JP02226.

PR 28-APR-1998; 98JP-0119395.

PA (SAGA) SAGAMI CHEM RES CENT.

PI (PROT-) PROTEGENE INC.

PT Kato S, Kimura T;

DR WPI: 2000-023358/02.

PS P-PSDB: AAY52391.

XX Claim 4; Page 106-107; 114pp; English.

CC This sequence represents the human cDNA clone HP10390  
 CC which encodes a 20 kD protein with one putative transmembrane  
 CC domain in the N-terminus. The cDNA was isolated from a human stomach  
 CC cancer cell line cDNA library. The protein has no homology with any  
 CC known protein. The protein may be used to raise specific antibodies, as  
 CC assay reagents, as diagnostic tissue markers, for the isolation of  
 CC cognate receptors, ligands and binding proteins, and as biologically  
 CC active agents. Nucleotides encoding the protein may be used as primers  
 CC and probes or antisense molecules, and in gene therapy. Cells transformed  
 CC with these nucleotides may be used to screen for agonists and antagonists  
 CC which are potentially useful therapeutically.

SO Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match 93.1%; Score 750.4; DB 21; Length 814;  
 Best Local Similarity 99.0%; Pred. No. 1,1e-179;

Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 11 CTGGGGCCAAAGTGAAGTCAAGCGCTGCGCAGCGCTTGGGCGACGCGCGCCCTGG 70  
 DB 52 CTGGGGCCAAAGTGAAGTCAAGCGCTTCCAGCGCTTGGGCGACGCGCGCGCCCTGG 111

QY	71	GACCAAGGTGGAGCAACCCCGTTACCTTAAATATGAAAGCTGGGGTTGGTGGCCCTG	130
Db	112	GAGCAGAGGTGGAGCAACCCCATTAACCTTAAAGATGAAAGCTGGGGTTGGTGGCCCTG	171
QY	131	CTTCTGGGGGCCCCCTGCTGGGAACCGCCTGGGCTCGAGAGACCAAGATCTCCACTGTGA	190
Db	172	CTTCTGGGGGCCCCCTGCTGGGAACCGCCTGGGCTCGGAGAGCAAGATCTCCACTGTGGA	231
QY	191	GCATGCAAGGCTCTGGTGGATGAACCTAGAAATGGGAATTTGCCAGGTGGACCCCAAGAA	250
Db	232	GCATGCAAGGCTCTGGTGGATGAACCTAGAAATGGGAATTTGCCAGGTGGACCCCAAGAA	291
QY	251	ACCAATTCAGATGGGATCTTTCCGGATCAATTCACAGATGGCAGCCAGTCAGTGGTGAAGTG	310
Db	292	ACCAATTCAGATGGGATCTTTCCGGATCAATTCAGATGGCAGCCAGTCAGTGGTGAAGTG	351
QY	311	CCCTATGCCCCCTCAGAGGGCCACCTCACAGACTGCTGGAGAGATATGTACCCGATG	370
Db	352	CCCTATGCCCCCTCAGAGGGCCACCTCACAGACTGCTGGAGAGATATGTACCCGATG	411
QY	371	AAGGAGTATGGGGGAACGATTGATCCTTCCACCACATGCAAGAACTTACGTACGTATAGTG	430
Db	412	AAGGAGTATGGGGGAACGATTGATCCTTCCACCACATGCAAGAACTTACGTACGTATAGTG	471
QY	431	GCCCGGAATGGAATTCACAGTGAACCTGGACCTTACAAAGCATTCGGAATCGACTCAGATATT	490
Db	472	GCCCGGAATGGAATTCACAGTGAACCTGGACCTTACAAAGCATTCGGAATCGACTCAGATATT	531
QY	491	AGCGGCAACCTCAAGTTTGGCTGGAGAGCATTTGTGGAGAAATACGAGATGAACCATTT	550
Db	532	AGCGGCAACCTCAAGTTTGGCTGGAGAGCATTTGTGGAGAAATACGAGATGAACCATTT	591
QY	551	GAATTCCTTTCCGAGAGGCTGACATGTTTAAAGACAAACTTTCGAGTAAGCAGACAGAT	610
Db	592	GAATTCCTTTCCGAGAGGCTGACATGTTTAAAGACAAACTTTCGAGTAAGCAGACAGAT	651
QY	611	CTTTGTGACCATGCGCCGACATTTGCGATGATGAGGTATGAACCATGTGAGAGCCGAC	670
Db	652	CTTTGTGACCATGCGCCGACATTTGCGATGATGAGGTATGAACCATGTGAGAGCCGAC	711
QY	671	ACTGGCTTGATGATCAACCCCAAGAGAGGGAATAATGGTGGCAATGCTTTATATATTAT	730
Db	712	ACTGGCTTGATGATCAACCCCAAGAGAGGGAATAATGGTGGCAATGCTTTATATATTAT	771
QY	731	GTTTTACTGAATTAACGTAAAAAATATATAAACCAAAAGTA	772
Db	772	GTTTTACTGAATTAACGTAAAAAATATATAAACCAAAAGTA	813

RESULT 4	
ABK52765	
ID	ABK52765 standard; cDNA; 814 BP.
XX	
AC	ABK52765;
XX	
DT	15-AUG-2002 (first entry)
XX	
DE	cDNA encoding transmembrane protein 4, a cancer-linked protein.
XX	
XX	Expressed sequence tag; EST; human; cancer; anti-neoplastic;
KM	cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12
KW	transmembrane protein 4; gene; ss.
XX	
OS	Homo sapiens.
XX	
EH	Location/Qualifiers
FT	145..693
CDS	
FT	/*tag= a
FT	/product= "Transmembrane protein 4"
XX	
PN	W0200231198-A2.
XX	

PD 18-APR-2002.  
XX  
PF 11-OCT-2001; 2001KO-US31607.  
XX  
PR 11-OCT-2000; 2000US-239294P.  
PR 11-OCT-2000; 2000US-239297P.  
PR 11-OCT-2000; 2000US-239605P.  
PR 12-OCT-2000; 2000US-239802P.  
PR 12-OCT-2000; 2000US-239805P.  
PR 12-OCT-2000; 2000US-239806P.  
PR 16-OCT-2000; 2000US-240622P.  
PR 19-OCT-2000; 2000US-241682P.  
PR 19-OCT-2000; 2000US-241723P.  
PR 31-OCT-2000; 2000US-244932P.  
XX  
PA (AVAL-) AVALON PHARM.

identifying modulators of a cancer-related gene to screen agents for preventing or treating cancer comprises detecting a difference in the expression of cancer-linked genes in the presence or absence of test compounds -

The invention relates to modulators of a cancer-related genes. Also described are: (1) processes for identifying an anti-neoplastic agents comprising contacting a cell exhibiting neoplastic activity with a compound first identified as a cancer related gene modulator, and detecting a decrease in the neoplastic activity. (2) a process for determining the cancerous state of a cell by determining an increase in the level of expression of at least one gene, where an elevated expression relative to a known non-cancerous cell indicates a cancerous state or potentially cancerous state. The anti-neoplastic agent is useful for treating cancer or for protecting an animal against cancer. The immunogenic composition is also useful for treating cancer in an animal, where the composition elicits the production of cytotoxic T lymphocytes specific for the immunogenic composition. Preferably, the animal is a human. The cancer-linked genes and polypeptides are also useful as targets for cancer therapy or chemotherapy. The present sequence represents a cancer-linked gene located on chromosome 12, which encodes transmembrane protein 4.

Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match	93.18;	Score 750.4;	DB 24;	Length 814;
Best Local Similarity	99.08;	Pred. No. 1.1e-179;		
Matches 754; Conservative	1;	Mismatches 7;	Indels 0;	Gaps 0;

Qy	11	CTGGGGCCAAAGTAAAGATGCACGGCTGACACGGTTGGGGCCACGGCGGCGCCCTGG	70
Db	52	CTGGGGCCAAAGTAAAGATGCACGGCTGACACGGTTGGGGCCACGGCGGCGCCCTGG	111
Qy	71	GACCAAAAGGTGAGCAACCCGTTACCCCTAAARATGAAAGGCTGGGCTTGCTGGCCCTG	130
Db	112	GAGCAAGGTGAGCAACCCGTTACCCCTAAAGATGAAAGGCTGGGCTTGCTGGCCCTG	177
Qy	131	CTTCTGGGGGCGCTGCGAGAACCGGCTGGGCTGCGAGAGACCGAGATCTCCACTGTGGA	199
Db	172	CTTCTGGGGGCGCTGCTGGGGAACCGGCTGGGCTGCGAGAGACCGAGATCTCCACTGTGGA	233
Qy	191	GCATGCAGGGCTCTGTGGATGAACCTAGAAATGGGAAATGGCCAGGTGCAACCCCAAG	250
Db	232	GCATGCAGGGCTCTGTGGATGAACCTAGAAATGGGAAATGGCCAGGTGCAACCCCAAG	291
Qy	251	ACCATTAGATGGGATCTTTCCCGGATCAATCCAGATGGCAGCAGTCAATGTGGAGGTG	310
Db	292	ACCATTAGATGGGATCTTTCCCGGATCAATCCAGATGGCAGCAGTCAATGTGGAGGTG	351

QY	311	CCCTATGGCCCTCAGAGGCCACCTCACAAGCTGCTGGAGGATATGTGACCGGAT	370
Db	352	CCCTATGGCCCTCAGAGGCCACCTCACAAGCTGCTGGAGGATATGTGACCGGAT	411
QY	371	AAGGAGTATGGGGAACAGATTGATCCTCCACCCTACGCAAGAATACGTAAGCTGTA	430
Db	412	AAGGAGTATGGGGAACAGATTGATCCTCCACCCTACGCAAGAATACGTAAGCTGTA	471
QY	431	GGCCGGAAATGGAGAAATCAGGAACTGGACCTCAAGGCAATCCGAATCGACTCA	490
Db	472	GGCCGGAAATGGAGAAATCAGGAACTGGACCTCAAGGCAATCCGAATCGACTCA	531
QY	491	AGCGGCAACCCCAAGTTTGGGTGAGAGCATTGTGGAGAAATACGAGATGAATCATT	550
Db	532	AGCGGCAACCCCAAGTTTGGGTGAGAGCATTGTGGAGAAATACGAGATGAATCATT	591
QY	551	GAATTCCTTTTCCGAGAGGCTGACATGTTAAAGCAAACCTTTCGATTAAGCAAC	610
Db	592	GAATTCCTTTTCCGAGAGGCTGACATGTTAAAGCAAACCTTTCGATTAAGCAAC	651
QY	611	CTTTGTGACATGCGCCGTCGACATATTCGATGAGTATGAACCACTGGAGAGGCCAC	670
Db	652	CTTTGTGACATGCGCCGTCGACATATTCGATGAGTATGAACCACTGGAGAGGCCAC	711
QY	671	ACTGGCTTGATGATCAACCCCGAGAGGGGAAAAATGCTGGCAATGCCCTTATATAT	730
Db	712	ACTGGCTTGATGATCAACCCCGAGAGGGGAAAAATGCTGGCAATGCCCTTATATAT	771
QY	731	GTTTTTACTGAATTAATCTGAAAAAATATGAACCAAAACTA	772
Db	772	GTTTTTACTGAATTAATCTGAAAAAATATGAACCAAAACTA	813

RESULT 5	
ABL41995	
ID	ABL41995 standard; DNA; 814 BP.
XX	
AC	ABL41995;
XX	
DT	11-JUN-2002 (first entry)
XX	
DE	Nucleotide sequence of human polypeptide HP10390.
XX	
XX	Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	CDS 145..693
FT	/*tag- a
FT	/product- "HP10390"
XX	
PN	WO200208416-A1.
XX	
PD	31-JAN-2002.
XX	
PF	24-JUL-2001; 2001WO-JP06371.
XX	
PR	24-JUL-2000; 2000JP-0222743.
PR	24-AUG-2000; 2000JP-0254407.
XX	
PA	(MISC-) JAPAN SCI & TECHNOLOGY CORP.
PI	Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;
XX	
DR	WPI; 2002-195877/25.
XX	
DR	P-PSDB; ABB09717.
XX	
PT	Antibody preparation by inoculation of an animal with a vector
PT	expressing a fusion protein of an antigen on the C-terminal side of a
PT	transmembrane domain for use as drugs, diagnostic reagents and
XX	laboratory reagents
XX	

PS Example, Page 33-35, 45pp. Japanese.  
XX The specification describes a method of antibody preparation. The  
CC method comprises inoculating an animal with a vector expressing  
CC a fusion protein having an antigen protein fused to the C-terminal  
CC side (extracellular) of a transmembrane domain protein (the  
CC N-terminal side of which is intracellular), and then isolating and  
CC purifying the antibody from the animal. The antibodies can be used  
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The  
CC present sequence encodes a polypeptide, designated HP10390, which was  
CC used in the course of the invention.

50 Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match	93.18; Score 750.4; DB 24; Length 814;
-------------	--

```

Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

QY 11 CTGGGCCCAAGTGAAAGTCCAGCGGCTCTGCCAGCGCTTGGGCCACGGCGGGCCCTGG 70

Db 52 CTGGGGCCAAAGTGAGAGTCCACGGGCTTCCACGGCCTGGGCCACGGCGGCCCTGG 11

QY 71 GACCAAGGTGGAGCAACCCGTTACCTAAARATGAAGGCTGGGGTTGGCTGGCCCTG 130

Db 112 GAGCAGAGTGGAGCGACCCATTACGCTAAAGATGAAGGCTGGGGTTGGCTGGCCCTG 177

QY 131 CTCTGGGGCCCTGCTGGGAACCGCCCTGGGCTCGGAGGAGCCAGGATCTCCACTGTGGA 190

Db 172 CTTCTGGGGCCCTGCTGGGAACCGCCCTGGGCTCGGAGGAGCCAGGATCTCCACTGTGGA 231

191 GCATGCAGGCTCTGGTGATGAATGGGAATTGCCAGGTGGACCCCAAG 250

Db 232 GCATGCAGGGCTCTGGTGGATGAAGTGGAAATGCCCGAGGTGGACCCCAAGAG 291

251 ACCATTGAGATGGATCTTCCGGATCAATCCAGATGGCAGCCAGTCAGTGGTGAGGTG 310

Db 292 ACCATTGATGGGATCTTCCGGATCAATCCAGATGGCAGCCAGTCAGTGCTGGAGGTG 351

311 CCTTATGCCGCTCAGAGGCCACCTCACAGAGCTGCTGGAGGAGATATGTGACCGGATG 370

Db 352 CCTATGCCCGCTCAGAGGCCACCTCACAGAGCTGCTGGAGGAGATATGTGACCGGATG 411

371 AAGGAGTATGGGACAGATTGATCCTTCCACCATCGCAGAACTACGTACGTAGTG 430

Db 412 AAGGATATGGGACAGATTGATCCTTCCACCCATCGCAAGACTACGTACGTAGTG 477

431 GCGCGGATGGAGATCCAGTGA CTGACCTACAAGGCATCCGAATCGACTCAGATATT 490

Db 472 GGCCGATGGAGATCCAGTGA CTGGACCTACAAGGCATCCGAATCGACTCAGATATT 533

491 AGCGCACCTCAAGTTTGCCTGTGAGAGCATTGTGAGGAATACGAGGATGAATCAT 550

Db 532 AGCGCACCTCAAGTTGCGTGTGAGAGCATTTGTGGAGGATACGAGGATGAATCATT 593

QY 551 GAATTCCTTTCCCGAGAGGCTGACAATGTTAAGACAACCTTGCAGTAAGCGACACAGAT 610

Db 592 GAATTCCTTTCCCGAGAGGCTGACAATGTTAAGACAACTTGCAGTAAGCGACAGAT 653

611 CTTTGACCATGCCCTGCACATATCGCATGATGAGCTATGAACTGAGACGCCCAC 670

Db 652 CTTGTGACCATGCCCTGCACATATCGCATGATGAGCTATGAACCACTGGAGCAGCCAC 711

QY 671 ACTGGCTTGATGATCACCCTCAGAGGGGAAATGGTGCATGCCCTTTATATATAT 730

Db 712 ACTGGCTTGATGATCACCCTCCAGGAGGGGAATGGTGGCAATGCCCTTTATATATAT 771

QY 731 GTTTTACTGAAATTACTGAAAAATATGAACCAAGTA 772

Db 772 GTTTTACTGAAATTAACTGAAAAATATGAACCAAAAGTA 813

## RESULTS

ABK09772

ID	ABK09772 standard; cDNA; 814 BP.
AC	ABK09772;
XX	
DT	14-MAR-2002 (first entry)
XX	
DE	Human ovarian tumour protein encoding cDNA #305.
XX	
KW	Human; ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;
KW	gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.
XX	
OS	Homo sapiens.
XX	
PN	WO200190154-A2.
XX	
PD	29-NOV-2001.
XX	
XX	23-MAY-2001; 2001WO-US16895.
XX	
PR	24-MAY-2000; 2000US-207107P.
PR	13-JUN-2000; 2000US-211457P.
PR	21-JUN-2000; 2000US-213673P.
PR	03-AUG-2000; 2000US-223288P.
PR	01-MAR-2001; 2001US-272190P.
XX	
XX	(CORI-) CORIXA CORP.
PA	
PI	Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MJ;
PI	Algate PA, Fling SP, Mannion J, Benson DR, Carter D;
DR	WPI; 2002-097641/13.
XX	
PT	New isolated polynucleotide encoding polypeptide comprising portion of
PT	ovarian tumour protein, useful for detection, diagnosis and therapy of
P1	human ovarian cancer
XX	
PS	Claim 1; Page 253-254; 285bp; English.
XX	
XX	The invention relates to an isolated polynucleotide encoding a
CC	polypeptide comprising a portion of an ovarian tumour protein. The
CC	sequences of the invention are useful for stimulating an immune response
CC	and for treating ovarian cancer in a patient. An antigen presenting cell
CC	that expresses the sequences is useful for treating ovarian cancer by
CC	incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells
CC	can then be proliferated and administered to the patient to inhibit the
CC	development of cancer. The DNA sequences are useful as probes or primers
CC	for nucleic acid hybridisation, to direct expression of a polypeptide in
CC	appropriate host cells. Detecting the presence of a cancer in a patient
CC	involves obtaining a biological sample from the patient, contacting the
CC	biological sample with an agent that binds to the protein, detecting the
CC	amount of protein that binds to the agent, comparing the amount of
CC	protein to a predetermined cut-off value and determining the presence of
CC	cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA
CC	molecules encoding ovarian tumour proteins of the invention.
XX	
XX	Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;
S0	
Query Match	93.1%; Score 750.4; DB 24; Length 814;
Query Similarity	99.0%; Pred. No. 1.1e-179;
Matches 754; Conservative 1; Mismatches 7; Indels 0; Gaps 0;	
11	CTGGGCGCCAAAGTGAAGTCCAGCGGTGCGCAGCGCTTGCGGCCAGCGGCGGCCCTGCG 70
Db	52 CTGGGCGCCAAAGTGAAGTCCAGCGGTGCTTCAGCGCTTGCGGCCAGCGGCGGCCCTGCG 111
0Y	71 GACCAAGGTGAGCAACCCCGTTACCTTAARATGAAAGCGTGGGGTTGCTGGCCCTG 130
Db	112 GAGCAGAGGTGAGCGACGCCCATTCAGCTAAGATGCAAAAGCGTGGGGTTGCTGGCCCTG 171
0Y	131 CTTTCGGGGGGCCCTCTCTCTGGGAAACCCCGTGGGTGCGAGAGAGCAGAGATCTCAGCTGTGA 190
Db	172 CTTTCGGGGGGCCCTCTCTCTGGGAAACCCCGTGGGTGCGAGAGAGCAGAGATCTCAGCTGTGA 231

QY	191	GGTCGAGGGGCTGCTGGTGGATGAACTGAAAGGGAATTTCCAGGTGGACCCCAAG	250
QY	191	GGTCGAGGGGCTGCTGGTGGATGAACTGAAAGGGAATTTCCAGGTGGACCCCAAG	250
Db	232	GGATGAGGGGCTCTGTGGATGAACCTGAAGGGGAAATTTCCAGGTGGACCCCAAG	291
QY	251	ACCATTTAGATGGGATCTTTTCCGATCAATCCAGATGGACCCAGTCAGTGTGAGTG	310
Db	292	ACCATTTAGATGGGATCTTTTCCGATCAATCCAGATGGACCCAGTCAGTGTGAGTG	351
QY	311	CCTTATGCCCCGTCAGAGGCCACCCTCAGAGAGCTGCTGGAGAGATATGTACCGGATG	370
Db	352	CCTTATGCCCCGTCAGAGGCCACCCTCAGAGAGCTGCTGGAGAGATATGTACCGGATG	411
QY	371	AAGGAGTATGGGGGAACAATTTGATTCCTTCCACCATTGCAAGAACTATAGTGTAGTG	430
Db	412	AAGGAGTATGGGGGAACAATTTGATTCCTTCCACCATTGCAAGAACTATAGTGTAGTG	471
QY	431	GGCCGGAATGAGATTCAGTGAATGGAACCTTACCAAGGCAATCCGAATTCAGTCAATTT	490
Db	472	GGCCGGAATGAGATTCAGTGAATGGAACCTTACCAAGGCAATCCGAATTCAGTCAATTT	531
QY	491	AGCGGACCCCTCAATTTGCTGCTGAGAGCAATTTGTGGAGGAATACGAGATCAATTT	550
Db	532	AGCGGACCCCTCAATTTGCTGCTGAGAGCAATTTGTGGAGGAATACGAGATCAATTT	591
QY	551	GAATTCCTTTTCCCGAGAGGCTGACAAATGTTTAAAGACAACTTTTGCATGTAAGCAGAT	610
Db	592	GAATTCCTTTTCCCGAGAGGCTGACAAATGTTTAAAGACAACTTTTGCATGTAAGCAGAT	651
QY	611	CTTTTGACCAATGCCCCCTGCACATTTTCGATGATGAGTATGAAACCACTGGAGAGCCAC	670
Db	652	CTTTTGACCAATGCCCCCTGCACATTTTCGATGATGAGTATGAAACCACTGGAGAGCCAC	711
QY	671	ACTGGCTTGATGATGATCACCCCCAGAGAGGGGAAAATGCTGGCAATGCTTTTATATATTT	730
Db	712	ACTGGCTTGATGATGATCACCCCCAGAGAGGGGAAAATGCTGGCAATGCTTTTATATATTT	771
QY	731	GTTTTTACTGGAATTTACTGAAAAAATATGAAACCAAGTA	772
Db	772	GTTTTTACTGGAATTTACTGAAAAAATATGAAACCAAGTA	813
RESULT 7			
AAK22112			
ID	AAK22112 standard; DNA; 1180 bp.		
XX	AAK22112;		
AC	AAK22112;		
DT	18-MAY-1999 (first entry)		
XX	Human secreted protein gene 2 clone H2MB56.		
DE	Human secreted protein gene 2 clone H2MB56.		
XX	Human; secreted protein; gene therapy; protein therapy; tissue; cancer;		
KW	tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;		
KW	developmental abnormality; foetal deficiency; Alzheimer's disease;		
KW	cognitive disorder; schizophrenia; immunological disorder; mood disorder		
KW	immune deficiency disease; respiratory disorder; arthritis; skeletal;		
KW	haematopoietic disorder; neural; osteoporosis; metabolic disorders;		
KW	cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	Homo sapiens.		
PN	W0901020-A2.		
XX	W0901020-A2.		
PD	14-JAN-1999.		
XX	14-JAN-1999.		
PF	30-JUN-1998; 98WO-0513608.		
XX	30-JUN-1998; 98WO-0513608.		
PR	12-SEP-1997; 97US-0058663.		
PR	01-JUL-1997; 97US-0051381.		
PR	01-JUL-1997; 97US-0051480.		
PR	12-SEP-1997; 97US-0058598.		
XX	12-SEP-1997; 97US-0058598.		
PA	(HUMA-) HUMAN GENOME SCI INC.		

```
XX Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM:
PI WPI: 1999-105683/09.
DR P-PSDB: AAY01136, AAY01162, AAY01163.
XX
PT New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, immune deficiency diseases or blood
PT disorders
XX
PS Claim 4; Page 117; 179pp: English.
XX
CC The invention relates to nucleic acid sequences (AA022111 to AA022134)
CC encoding human secreted proteins (AAY01135 to AAY01158). The secreted
CC protein gene sequences are deposited with the ATCC under deposit number
CC ATCC 209118. Host cells comprising recombinant vectors containing the
CC nucleic acid sequences are used for the recombinant production of the
CC secreted proteins. The polynucleotide and amino acid sequences are useful
CC for are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Pathological conditions can
CC be also diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, developmental abnormalities and foetal deficiencies,
CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,
CC schizophrenia, immunological disorders, immune deficiency diseases
CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,
CC haematopoietic disorders, neural disorders, skeletal disorders,
CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine
CC disorders or gastrointestinal disorders. The polypeptides are also useful
CC for identifying their binding partners. The present sequence represents a
CC gene encoding a human secreted protein (see descriptor line for gene
CC number and clone identification).
XX
SQ Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other:
XX
Query Match 93.0%; Score 749.4; DB 20; Length 1180;
Best Local Similarity 98.9%; Pred. No. 2,2e-179;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
OY 11 CTGGGGCCAAAGTGAAGTCCAGCGGTCTGCGAGCGCTGGGCGCGCGCCCTGG 70
DB 420 CTGGGGCCAAAGTGAAGTCCAGCGGTCTGCGAGCGCTGGGCGCGCGCCCTGG 479
OY 71 GACCAAAAGTGGAGCAACCCGTTACCTTAATAATGAAGGCTGGGGTTGGCTGGCCTG 130
DB 480 GACCAAGGTGGAGCAACCCGTTACCTTAATAATGAAGGCTGGGGTTGGCTGGCCTG 539
OY 131 CTCTGGGGGCGCTGCGCAACCGCTGGGCTCGAGAGCGCAGAGATCCACTGTGGA 190
DB 540 CTCTGGGGGCGCTGCTGGGAACCGCTGGGCTCGAGAGCGCAGAGATCCACTGTGGA 599
OY 191 GCATGAGGGCTGTGTGATGAGTAAATGGAAATTCGCCAGGTGGAAGAG 250
DB 600 GCATGAGGGCTGTGTGATGAGTAAATGGAAATTCGCCAGGTGGAAGAG 659
OY 251 ACCATTCAGTGGGATCTTCCGGATCAATCCAGATGGCGCAGTCACTGAGAGGTG 310
DB 660 ACCATTCAGTGGGATCTTCCGGATCAATCCAGATGGCGCAGTCACTGAGAGGTG 719
OY 311 CTTTATGCCCGCTCAGAGCGCCACTCAGAGCTGCTGAGAGAGATATGTGACGGATG 370
DB 720 CTTTATGCCCGCTCAGAGCGCCACTCAGAGCTGCTGAGAGAGATATGTGACGGATG 779
OY 371 AAGGAGTAGGGGAACAGATGTATCTTCCAGCCATCGAAGAACTACGTGTATGAGT 430
DB 780 AAGGAGTAGGGGAACAGATGTATCTTCCAGCCATCGAAGAACTACGTGTATGAGT 839
OY 431 GGCGGGAATGAGAAATCCATGAACTGAGACTCAAGGATCCGAATCGACTCAGATAT 490
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

```
DB 840 GGCGGGAATGAGAAATCCATGAACTGAGACTCAAGAGCAATCCGAATCGACTCAGATAT 899
OY 491 AGCGGACCCCTCAAGTGTTCGTGTGAGAGCAATGTGAGAGAAATACGAGATCAATCAT 550
DB 900 AGCGGACCCCTCAAGTGTTCGTGTGAGAGCAATGTGAGAGAAATACGAGATCAATCAT 959
OY 551 GAATTTTTCCTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAAGCAAGAT 610
DB 960 GAATTTTTCCTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACGTAAAGCAAGAT 1019
OY 611 CTTTGTGACCAACCCCTCGACATATGCAATGAGTACATGCAATGAGAGAGAGAGAGAG 670
DB 1020 CTTTGTGACCAACCCCTCGACATATGCAATGAGTACATGCAATGAGAGAGAGAGAGAG 1079
OY 671 ACTGCGTATGATGATCAACCCCGAGAGGGAATAATGTCGCAATGCTTTATATATATAT 730
DB 1080 ACTGCGTATGATGATCAACCCCGAGAGGGAATAATGTCGCAATGCTTTATATATATAT 1139
OY 731 GTTTTACTGAAATTAATGTAAGTAAATATGAAACCAAAAGT 771
DB 1140 GTTTTACTGAAATTAATGTAAGTAAATATGAAACCAAAAGT 1180
RESULT 8
AAS21330
ID AAS21330 standard; cDNA; 1210 BP.
AC AAS21330;
XX
AC AAS21330;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human cDNA sequence encoding for PRO4426 polypeptide.
XX
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW breast; prostate; cervical; tumour necrosis factor- $\alpha$ ; TNF- $\alpha$ phn;
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200140466-A2.
XX
PD 07-JUN-2001.
XX
OY 01-DEC-2000; 2000MO-US32678.
XX
PR 01-DEC-1999; 99MO-US28301.
PR 01-DEC-1999; 99MO-US28634.
PR 02-DEC-1999; 99MO-US28551.
PR 02-DEC-1999; 99MO-US28564.
PR 02-DEC-1999; 99MO-US28565.
PR 09-DEC-1999; 99US-0170262.
PR 16-DEC-1999; 99MO-US30095.
PR 20-DEC-1999; 99MO-US30911.
PR 20-DEC-1999; 99MO-US30999.
PR 30-DEC-1999; 99MO-US31243.
PR 06-JAN-2000; 2000MO-US00277.
PR 06-JAN-2000; 2000MO-US00376.
PR 11-FEB-2000; 2000MO-US03565.
PR 18-FEB-2000; 2000MO-US04341.
PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
PR 24-FEB-2000; 2000MO-US04914.
PR 24-FEB-2000; 2000MO-US05004.
PR 01-MAR-2000; 2000MO-US05601.
PR 20-MAR-2000; 2000MO-US07377.
PR 21-MAR-2000; 2000MO-US07332.
PR 30-MAR-2000; 2000MO-US08439.
PR 17-MAY-2000; 2000MO-US13705.
PR 22-MAY-2000; 2000MO-US14042.
PR 30-MAY-2000; 2000MO-US14941.
PR 02-JUN-2000; 2000MO-US15264.
PR 10-NOV-2000; 2000MO-US30873.
```

XX (GENE ) GENE/TECH INC.  
 XX Baker KP, Beresini M, DeForge L, Desnoyers L, Flamaroff E, Gao W;  
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;  
 XX WPI: 2001-408281/43.  
 DR P-PSDB: AAU12258.  
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX Claim 3; Fig 173; 813pp; English.  
 PS  
 XX AA521244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX  
 SO Sequence 1210 BP: 261 A: 316 C: 379 G: 254 T: 0 other:  
 Query Match 93.0%; Score 749.4; DB 22; Length 1210;  
 Best Local Similarity 98.9%; Pred. No. 2.2e-179;  
 Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

DB 870 GCGCGAATGAGAAATCCAGTCACTGACCTACAGAGCATCCGATGACCTCAGATATT 929  
 OY 491 AGCGGACACCTCAAGTTTCGTGTGAGAGACATTGTGAGAGAAATAGAGATCACTATT 550  
 DB 930 ACGGCGACCTCAAGTTTCGTGTGAGAGACATTGTGAGAGAAATAGAGATCACTATT 989  
 OY 551 GAATTCCTTTCCCGAGAGGCTGACATATGTTAAAGACAACTTTGCAATGAGCAACAGAT 610  
 DB 990 GAATTCCTTTCCCGAGAGGCTGACATATGTTAAAGACAACTTTGCAATGAGCAACAGAT 1049  
 OY 611 CTTTGTGACCATGACCTGACATATGTCATGATGAGCTGATGAACCACTGAGACAGCCAC 670  
 DB 1050 CTTTGTGACCATGACCTGACATATGTCATGATGAGCTGATGAACCACTGAGACAGCCAC 1109  
 OY 671 ACTGCTTGATGATGATCACCAGGAGGAGGAAATGTTGCAATGCTTTATATATAT 730  
 DB 1110 ACTGCTTGATGATGATCACCAGGAGGAGGAAATGTTGCAATGCTTTATATATAT 1169  
 OY 731 GTTTTACTGAAATTAACGAAAAAATATGAAACCAAAAGT 771  
 DB 1170 GTTTTACTGAAATTAACGAAAAAATATGAAACCAAAAGT 1210  
 RESULT 9  
 ID AAX97884 standard; cDNA: 714 BP.  
 XX AAX97884;  
 XX 23-SEP-1999 (first entry)  
 DE Human secreted protein encoding cDNA #72.  
 XX Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
 KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
 XX Homo sapiens.  
 OS  
 XX W09925825-A2.  
 XX 27-MAY-1999.  
 XX 13-NOV-1998; 98WO-1B01862.  
 XX 04-SEP-1998; 98US-0099273.  
 XX 13-NOV-1997; 97US-0066677.  
 XX 17-DEC-1997; 97US-0066957.  
 XX 09-FEB-1998; 98US-0074121.  
 XX 13-APR-1998; 98US-0081563.  
 XX 10-AUG-1998; 98US-0096110.  
 XX (GIST ) GENESET.  
 PA Bougueleret L, Duclert A, Dumas Milne Edwards J;  
 XX WPI: 1999-347472/29.  
 DR P-PSDB: AAY36200.  
 XX Extended cDNAs encoding secreted proteins  
 PT Claim 1; Page 272; 307pp; English.  
 PS  
 XX AAY97813-X97906 represent extended cDNA's which encode novel human  
 CC secreted proteins (see AAY36129-X36222) and which have cytosolic,  
 CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
 CC express secreted proteins or parts of them or to obtain antibodies  
 CC capable of binding to the secreted proteins. They may also be used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC Uses also include design of expression vectors and secretion vectors.  
 XX  
 SO Sequence 714 BP: 203 A: 158 C: 202 G: 151 T: 0 other:



Query Match 87.1%; Score 702.4; DB 20; Length 714;  
Best Local Similarity 98.9%; Pred. No. 1.3e-167;  
Matches 706; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 72 ACCAAGGAGGAGCAACCCGTTACCTTAARATGAAGAGCGGGTGGCGCCCTGC 131  
DB 1 ACCAGAGGAGGAGCGACCCATTACCTTAAGATGAAGAGCGGGTGGCGCCCTGC 60  
QY 133 TTCTGGGGCCCTGCTGGGAGACCGCTGGGCTGCGAGAGAGATGCCACTGTGAG 191  
DB 61 TTCTGGGGCCCTGCTGGGAGACCGCTGGGCTGCGAGAGAGATGCCACTGTGAG 120  
QY 192 CATGCAAGGCTCTGTGGTGAATGAATGAATGGAAATGCCAGGTGAGACCCAGACA 251  
DB 121 CATGCAAGGCTCTGTGGTGAATGAATGAATGGAAATGCCAGGTGAGACCCAGACA 180  
QY 252 CCATTCAATGGAGATCTTCGCGATCAATCCAGATGGCCAGTCAGTGGTGGAGTGC 311  
DB 181 CCATTCAATGGAGATCTTCGCGATCAATCCAGATGGCCAGTCAGTGGTGGAGTGC 240  
QY 312 CTATGCCCCGCTCAGAGGCCCACTCACAGAGCTGCTGGAGAGATATGTGACCGGATG 371  
DB 241 CTATGCCCCGCTCAGAGGCCCACTCACAGAGCTGCTGGAGAGATATGTGACCGGATG 300  
QY 372 AGGAGTATGGGAGACAGATTGATCTCCACCACATCGCAAGAACTAGCTGTAGTG 431  
DB 301 AGGAGTATGGGAGACAGATTGATCTCCACCACATCGCAAGAACTAGCTGTAGTG 360  
QY 432 GCGGAGTATGGAGATTCAGTGAAGTGCATCAAGGAGTCCGAATCGAGCTAGATTA 491  
DB 361 GCGGAGTATGGAGATTCAGTGAAGTGCATCAAGGAGTCCGAATCGAGCTAGATTA 420  
QY 492 GCGGAGTATGGAGATTCAGTGAAGTGCATCAAGGAGTCCGAATCGAGCTAGATTA 551  
DB 421 GCGGAGTATGGAGATTCAGTGAAGTGCATCAAGGAGTCCGAATCGAGCTAGATTA 480  
QY 552 AATTCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACATAAGCGACAGATC 611  
DB 481 AATTCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTTGACATAAGCGACAGATC 540  
QY 612 TTGTGACCATGCGCTGCATATGATGATGATGATGATGATGATGATGATGATGATG 671  
DB 541 TTGTGACCATGCGCTGCATATGATGATGATGATGATGATGATGATGATGATGATG 600  
QY 672 CTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 731  
DB 601 CTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
QY 732 TTTTACTGATAATTAATGATAAATATGATAAATGATAAATGATAAATGATAAATG 785  
DB 661 TTTTACTGATAATTAATGATAAATATGATAAATGATAAATGATAAATGATAAATG 714

RESULT 10  
AAK97837  
ID AAK97837 standard; cDNA; 832 BP.  
XX AAK97837;  
AC AAK97837;  
XX 23-SEP-1999 (first entry)  
DT  
XX Human secreted protein encoding cDNA #25.  
DE  
XX Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX Homo sapiens.  
OS  
XX W09925825-A2.  
PN  
XX 27-MAY-1999.  
PD  
XX 13-NOV-1998; 98WO-1B01862.

XX 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.

XX (GEST) GENSET.  
XX  
XX Bouquelier L, Duclert A, Dumas Milne Edwards J;  
XX  
XX WPI: 1999-347472/29.  
XX P-PSDB: AAY36153.  
XX  
XX Extended cDNAs encoding secreted proteins  
XX  
XX Example 28; Page 194-195; 307p; English.  
XX  
XX AAY97813-X97906 represent extended cDNA's which encode novel human  
XX secreted proteins (see AAY36129-Y36222) and which have cytosolic,  
XX thrombotic and osteopathic activity. The extended cDNAs can be used to  
XX express secreted proteins or parts of them or to obtain antibodies  
XX capable of binding to the secreted proteins. They may also be used in  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX Uses also include design of expression vectors and secretion vectors.

Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other;

Query Match 85.4%; Score 688.4; DB 20; Length 832;  
Best Local Similarity 97.6%; Pred. No. 4.9e-164;  
Matches 761; Conservative 1; Mismatches 12; Indels 6; Gaps 6;

QY 11 CTGGGGCCAAAGAGAAAGTCCAGGCTGCGGCGCTGGGCGAGCGGCGCCCTGG 70  
DB 55 CTGGGGCCAAAGTGAAGATCCAGGCTGCGGCGCTGGGCGAGCGGCGCCCTGG 114  
QY 71 GACCAAGGTGGAGCAACCCCGTTACCTTAARATGAAGAGGCTGGGCTGGGCGCTG 130  
DB 115 GACCAAGGTGGAGCAACCCCGTTACCTTAARATGAAGAGGCTGGGCTGGGCGCTG 174  
QY 131 CTCTCTGGGGCCCTGCTGGGAGCCGCTGGGCTGGGAGAGCCAGATCTCCACTGTGA 190  
DB 175 CTCTCTGGGGCCCTGCTGGGAGCCGCTGGGCTGGGAGAGCCAGATCTCCACTGTGA 234  
QY 191 GCATGAGGGCTGTGGTGTGATG-ACATGATGGGAATTTGCCAGGTGAGACCCCAAGA 249  
DB 235 GCATGAGGGCTGTGGTGTGATG-ACATGATGGGAATTTGCCAGGTGAGACCCCAAGA 294  
QY 250 GACCATTCAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGTCAAGTGTGAGGT 309  
DB 295 GACCATTCAGATGGATCTTTCCGATCAATCCAGATGGCAGCAGTCAAGTGTGAGGT 354  
QY 310 GCCTTATGCCCGCTCAGAGGCCCACTCACAGAGCTGCTGGAGAGATATGACCGGAT 369  
DB 355 GCCTTATGCCCGCTCAGAGGCCCACTCACAGAGCTGCTGGAGAGATATGACCGGAT 414  
QY 370 GAAGAGTATGGGAGACAGATTGATCTCCAGCCTTCGCAAGAACTAGCTAGCTAGT 429  
DB 415 GAAGAGTATGGGAGACAGATTGATCTCCAGCCTTCGCAAGAACTAGCTAGCTAGT 474  
QY 430 GGGCCGGAATGGAATCCAGTGAAGTCACTCAAGAGCATCCGATCACTCACTATAT 489  
DB 475 GGGCCGGAATGGAATCCAGTGAAGTCACTCAAGAGCATCCGATCACTCACTATAT 534  
QY 490 TACGGCAGCCCTTAAG-TTGGCTGTGAGAGCATTTGTGAGAGAAATGAGAGATGAATCA 548  
DB 535 TACGGCAGCCCTTAAG-TTGGCTGTGAGAGCATTTGTGAGAGAAATGAGAGATGAATCA 594  
QY 549 TTGAATTTCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTGCAGTAACGGAACG 608  
DB 595 TTGAATTTCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTGCAGTAACGGAACG 654





PN W09901554-A1.  
 XX 14-JAN-1999.  
 XX 02-JUL-1998: 98MO-US13859.  
 XX 17-JUN-1998: 98US-0099005.  
 PR 03-JUL-1997: 97US-0051704.  
 PR 03-JUL-1997: 97US-0888088.  
 PR 19-MAY-1998: 98US-0081338.  
 PR 19-MAY-1998: 98US-0085983.  
 PR 17-JUN-1998: 98US-0089899.  
 XX (ZYMO ) ZYMOGENETICS INC.  
 PA Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;  
 PI WPI: 1999-106055/09.  
 DR P-PSDB; AAM88476.  
 DR  
 XX  
 XX  
 PT New mammalian secretory peptide-9 (Zs19) - used as a growth  
 PT enhancer for placenta, liver and heart, and as an indicator of  
 PT cancer  
 PS Claim 2: page 75-77; 85pp; English.  
 XX  
 XX This cDNA clone encodes novel mouse secretory peptide-9, or Zs19  
 CC (see AAM88476), an orthologue of novel human Zs19 (see AAM88469).  
 CC Human Zs19 is overexpressed in a number of tumours including  
 CC brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid  
 CC and lymphoma tumors, and thus can be used as an indicator for  
 CC cancer. The invention provides polynucleotides (see AAX06968-70)  
 CC encoding Zs19 polypeptides (see AAM88469-77) including mature  
 CC polypeptides, other processed forms, variants and the mouse  
 CC orthologues. The Zs19 gene, or probes derived from it, can be  
 CC used to determine if Zs19 is present on chromosome 10, and if a  
 CC mutation has occurred. Antibodies raised against Zs19 can be  
 CC used as diagnostic agents to determine the presence of Zs19, and  
 CC thus the presence of cancer. They can also be labelled with  
 CC radioisotopes or fused with toxins and used to treat tumours  
 CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
 CC cDNA can also be used to inhibit the growth of tumor cells. Zs19  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.  
 CC  
 XX  
 SQ Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;  
 Query Match 74.3%; Score 599; DB 20; Length 1069;  
 Best Local Similarity 87.7%; Pred. No. 2,1e-141;  
 Matches 690; Conservative 1; Mismatches 86; Indels 10; Gaps 3;

Db 572 CCCGCTCAGAGCCCACTTCACAGAGTTCCTTGAGAGAGGTGTGACCCGAATGAAGAGT 631  
 Qy 378 ATGGGGAAGATATGATCTTCCACCACCATCGCAAGAACTAGTGTAGTGGCCGGA 437  
 Db 632 ACGGGGAAGATTTGACCTTCTTACCCACCGCAAGAACTAGTGTGTGAGCCGGA 691  
 Qy 438 ATGAGAGATTCAGTGAACCTGACCTTACCAAGCATCCGAATGCAGATATTAGCCGGA 497  
 Db 692 ATGAGAGATTCAGTGAACCTTACAGGAGCATCCGAATGACTCAGATATCAGCCGGA 751  
 Qy 498 CCTCAAGTTGGCTGTGAGAGCATTTGTGAGAGAAATCGAGATGAACTATGAAATCT 557  
 Db 752 CCTCAAGTTGGCTGTGAGAGCATTTGTGAGAGAAATCGAGATGAACTATGAAATCT 811  
 Qy 558 TTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTCAGTAAGCGAAAGATCTTGTG 617  
 Db 812 TCTCCAGAGAGGCTGACAAATGTTAAAGCAAACTTTCAGTAAGCGAAAGATCTTGTG 871  
 Qy 618 ACATGCGCTGCAATATGCGATGATGAGTATGAAACCTGAGACGCCAC- -AC 672  
 Db 872 ACCATGCGCTGCAATATGCGATGATGAGTATGAAACCTGAGACGCCACCTACACC 931  
 Qy 673 TGGCTTATGATGATCAACCCCGAGAGGGGAAATGCTGCAATGCTTTATATATTATGT 732  
 Db 932 AAAGTGATGATGATCAACCCCGAGAGGGGAAATGCTGCAATGCTTTATATATTATGT 989  
 Qy 733 TTTTACTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789  
 Db 990 TTTTATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1049  
 Qy 790 AGAGAGA 796  
 Db 1050 AAAAAA 1056  
 RESULT 13  
 AAZ08294  
 ID AAZ08294 standard; DNA; 1069 BP.  
 AC AAZ08294;  
 XX  
 DT 07-FEB-2000 (first entry)  
 XX  
 DE Mouse ortholog gene encoding Zs19 secretory protein.  
 XX  
 KM Secretory protein-9; mouse Zs19; ortholog; overexpression; antagonist;  
 KM antibody; antisense nucleotide; tumour; treatment; receptor;  
 KM radio-label; polypeptide toxin; down-regulation; diagnostic;  
 KM therapeutic; probe; cancer; brain; liver; detection; stomach;  
 KM lymphoma; ds.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FH CDS 358..906  
 FT /\*tag- a  
 FT /product- "Mouse Zs19 secretory protein"  
 FT /note- "Overexpressed in tumours"  
 FT 358..417  
 FT /\*tag- b  
 FT mat\_peptide 418..903  
 FT /\*tag- c  
 FT /label- Mature\_Zs19\_protein\_variant-4  
 PN W09960405-A1.  
 PD 25-NOV-1999.  
 XX  
 XX 19-MAY-1999: 99MO-US11107.  
 PF 19-MAY-1998: 98US-0081183.  
 PR 19-MAY-1998: 98US-0081183.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.

xx Moore EE, Taft DM;  
xx  
xx MPI: 2000-039447/03.  
DR P-P8DB; AA1Y15136.  
xx  
xx  
xx Detecting tumors using antibodies, antagonists and antisense  
nucleotides to secretory protein-9 (Zs19) -  
PS Disclosure; Page 37-38; 45pp; English.

The present DNA sequence is an ortholog encoding the secretory protein-9,  
Zs19 derived from mouse. It is overexpressed in tumors. Antagonists,  
antibodies and antisense nucleotides to Zs19 are useful for detecting  
and treating tumors. The antagonist may be an antibody or receptor to  
Zs19 and it may be radio-labelled or fused to a polypeptide toxin. It  
can be used for down regulating the overexpression of Zs19. The gene  
sequence can be used as nuclear acid probes to detect RNA encoding Zs19.  
The Zs19 sequence facilitates improved diagnostic and therapeutic  
techniques for detecting and treating cancers, especially of the brain,  
liver, stomach, lymphoma etc., at an early stage.

Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other:

Query Match 74.3%; Score 599; DB 21; Length 1069;  
Best Local Similarity 87.7%; Pred. No. 2.1e-141;  
Matches 690; Conservative 1; Mismatches 86; Indels 10; Gaps 3

18 CAAAGTGAAGAAGCCACGCGGTCTGCCAGCCTTGCGGCACGGCGGCGCCTGGACCATAA 77  
Db CAAGATGAGAGTCGCGGTCTGCCAGACCTGGCGCACGGCGGCGCCTGGAGACA 331  
  
78 GGTGACGACCAACCCGTTACCTTAABAATGAAGAAGCTGGGCTGGCCCTGCTTGTG 137  
Db GGATGAGGAGAACCTTTACTTAAGATGAAGAAGCTGGGCTGGCTTGTG 391  
  
138 GGGCCCTGCTGGGAACCGCGCTGGCGGAGAGAGCAGATCTCCACCTGTGGAGATGCA 197  
Db GGGCTCTGCTGGGAACCTGCTGGCTCGAAGAGAGCCAGATTTACACTGTGAGCTTGA 451  
  
198 GGGCTCTGCTGGATTAAGATGGAATTGCCAAGGTGAGACCCCAAGAAGACCATTC 257  
Db GGGCTCTGCTGGATTAAGATGGAATTGCCCGCTGGACCCCAAGAAGACCATTC 511  
  
258 AGATGGGATCTTTCCGGGANTCATCCAGATGGAGCAGCAGTGTGGAGGCGCTTAG 317  
Db AGATGGGATCTTTCCGGGANTCATCCAGATGGAGCAGCAGTGTGGAGGCGCTTAG 571  
  
318 CCCGCTCAGAGGCCACACTCACAGAGCTGCTGGAGAGATATGTACCCGATGAAGAGAT 377  
Db CCCGCTCAGAGGCCACACTCACAGAGTGTCTGAGAGAGTGTGTACCGAATGAAGAGAT 631  
  
378 ATGGGGAACAGATTATCTTCCACCACATCGCAAGACTACGTAGTGTGGGCCGGA 437  
Db ACGGGGAACAGATTATCTTCCACCACCGCAAGACTACGTAGTGTGGGCCGGA 691  
  
438 ATGGAGATCCAGTAACTGAGACATCAAGAGGATCGAATGACATCAGATATTAGGGGA 497  
Db ATGGAGATCCAGTAACTGAGACATCAAGAGGATCGAATGACATCAGATATTAGGGGA 751  
  
632 ATGGAGATCCAGTAACTGAGACATCAAGAGGATCGAATGACATCAGATATTAGGGGA 751  
498 CCTTCAGATTGCGTGTGAGAGCATTTGGAGGAATACGAGGATCAATCATTTGAATCT 557  
Db CCTTCAGATTGCGTGTGAGAGCATTTGGAGGAATACGAGGATCGAATCATTTGAATCT 811  
  
558 TTTCOCGAGAGCTGACAAATGTTAAAGCAAACCTTGGCATGAAGGAACAGATCTTTGTG 617  
Db TTTCOCGAGAGCTGACAAATGTTAAAGCAAACCTTGGCATGAAGGAACAGATCTTTATGTG 871  
  
618 ACAATGCCCTGCATATATGCGATGATGAGAGCTGTGAACACATGGACAGAGCCAC-----AC 672  
Db ACAATGCCCTGCATATATGCGATGATGAGAGCTGTGAATCATCTGAGACAGACCATCAGACC 931  
  
673 TTGGCTTGAATGCAATCAACCCCCACAGAGGGAATAATGGTGGCAATGCTTTTATATATATGT 732

Db	932	AAACGTGATGGGAACACCCTCCAGAGGGGAAGATGCAGCATTCCTT--TATTATACGT	989
Oy	733	TTTTACTGAAATAAATCTGAAAAA---TATGAACAACCAAAAGTAATAAAAAAAAAAAAAAAAG	789
Dd	990	TTTTATGCAAAAGAAGACTGAAAAAACCTTGAAACCGAAGTAGTAAAAAAAAAAAAAAAAAAAA	1049
Oy	790	AGAGAGA	796
Db	1050	AAAAAAA	1056
RESULT 14 AAZ38326 ID AAZ38326 standard; cDNA; 546 BP. XX AC XX AAZ38326; XX DT 09-FEB-2000 (first entry) XX DE Human transmembrane protein cDNA clone HP10390 coding sequence. XX KW HP10390; transmembrane domain; stomach cancer cell; antibody; RV assay reagent; diagnostic marker; primach probe; antisense; gene therapy; KW agonist; antagonist; ligand; therapeutic; ds. XX OS Homo sapiens. FH FX Key Location/Qualifiers FT CDS 1..540 FT /*tag= a FT /product= "Human transmembrane protein HP10390" FT /note= "No stop codon given in the specification"FN WO9955862-A2. PN PD 04-NOV-1999. PP 27-APR-1999; 99WO-JP02226. PR 28-APR-1998; 98JP-0119395. PA (SAGA ) SAGAMI CHEM RES CENT. PA (PROT-) PROTEGENE INC. PI Kato S, Kimura T; DR WPI: 2000-023358/02. DR P-PSTDB; MAY52391. PT Human proteins with transmembrane domains, involved in control of cell proliferation and differentiation, useful for treating e.g. cancer or inflammation - XX XX Claim 3; Page 88; 114pp; English.			
Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;			
Query Match	67.7%; Score 546; DB 21; Length 546;		
Best Local Similarity	100.0%; Pred. NO. 3,9e-128;		
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

OY 104 ATGAAAGCTGGGTTGGCTGGCCCTGCTTGGGGGCCCTGCTGGGAACCGCCTGGGCT 163  
 DB 1 ATGAAAGCTGGGTTGGCTGGCCCTGCTTGGGGGCCCTGCTGGGAACCGCCTGGGCT 60  
 OY 164 CGGAGAGCCAGATCTCCACTGTGAGCATGCAAGGCTCTGGTGAATGAACTAGATG 223  
 DB 61 CGGAGAGCCAGATCTCCACTGTGAGCATGCAAGGCTCTGGTGAATGAACTAGATG 120  
 OY 224 GAAATGGCCAGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGATCATTCCA 283  
 DB 121 GAAATGGCCAGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGATCATTCCA 180  
 OY 284 GATGAGCAGCAGTCACTGTGTGAGTGCCCTTATGCCCGCTCAGAGGCCACCTCAGAG 343  
 DB 181 GATGAGCAGCAGTCACTGTGTGAGTGCCCTTATGCCCGCTCAGAGGCCACCTCAGAG 240  
 OY 344 CTGCTGGAGAGATATGTGACCCGATGAAGAGATATGGGGAACAGATTGCTTCCACC 403  
 DB 241 CTGCTGGAGAGATATGTGACCCGATGAAGAGATATGGGGAACAGATTGCTTCCACC 300  
 OY 404 CATCGAAGAACTACGTACGTAGTGGGCCGGAATGGAATCAGTGAATGGAACCTA 463  
 DB 301 CATCGAAGAACTACGTACGTAGTGGGCCGGAATGGAATCAGTGAATGGAACCTA 360  
 OY 464 CAAGCATCCGAATCGACTGAGATATTTAGCGGACCCCTCAAGTTTGGCTGAGACATT 523  
 DB 361 CAAGCATCCGAATCGACTGAGATATTTAGCGGACCCCTCAAGTTTGGCTGAGACATT 420  
 OY 524 GTGAGGAATACGAGATGAACTGATGATCTTTCCCGAGAGGCTGACATGTTAAA 583  
 DB 421 GTGAGGAATACGAGATGAACTGATGATCTTTCCCGAGAGGCTGACATGTTAAA 480  
 OY 584 GACAACTTTCAGTAAAGCAAGACATCTTTGACCATGCCCTGACATATGATGAT 643  
 DB 481 GACAACTTTCAGTAAAGCAAGACATCTTTGACCATGCCCTGACATATGATGAT 540  
 OY 644 GAGCTA 649  
 DB 541 GAGCTA 546  
 RESULT 15  
 AAA43525  
 ID AAA43525 standard; cDNA; 592 BP.  
 AC AAA43525;  
 XX  
 DT 21-AUG-2000 (first entry)  
 DE Mouse secreted expressed sequence tag SEQ ID NO:100.  
 XX  
 XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
 KW expressed sequence tag; EST; probe; chemokine; analgesic; haemostatic;  
 KW immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytoskeletal; antibacterial; antifungal;  
 KW antiviral; antidiabetic; antischistosomal; vulnery; antiparkinsonian;  
 KW anticancer; osteoprotective; neuroprotective; nootropic; antiparasitic;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
 KW tumour; infection; depression; psoriasis; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN W0200021991-A1.  
 XX  
 PD 20-APR-2000.

PF 15-OCT-1999; 99WO-US24206.  
 XX  
 PR 15-OCT-1998; 98US-0104436.  
 XX  
 PA (GENE) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR;  
 XX  
 DR MPI: 2000-317938/27.  
 XX  
 PT Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (SESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 XX  
 PS Claim 1: Page 228-229; 803pp; English.  
 XX  
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (SESTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The SESTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytoskeletal; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antischistosomal; vulnery; antiparkinsonian; antiparasitic;  
 CC neurotropic; antiparkinsonian; antiparasitic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
 CC therapy and in vaccines. The SESTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 CC  
 CC  
 CC  
 SO Sequence 592 BP; 159 A; 139 C; 167 G; 123 T; 4 other;  
 Query Match 54.5%; Score 439.2; DB 21; Length 592;  
 Best Local Similarity 89.8%; Pred. No. 3; 8e-101;  
 Matches 501; Conservative 1; Mismatches 53; Indels 3; Gaps 3;  
 OY 80 TGGAGCAACCCGTTACCTAAATGAAGGCTGGGCTGGTGG-CCTGCTTGGG 138  
 DB 20 TAGAGCAGCCCTGTACATGAAGATGAAGGTTGGTGGCTAGCCCTACTTGGG 79  
 OY 139 GGCCTGCTGGGAACCCGCTGGGCTCGAGAGAGCAGATCTCCACTGTGAGCATG 198  
 DB 80 GGTCTGCTGGGAACCCGCTGGGCTCGAGAGAGCAGATCTCCACTGTGAGCATG 139  
 OY 199 GGCCTGCTGGGAACCCGCTGGGCTCGAGAGAGCAGATCTCCACTGTGAGCATG 258  
 DB 140 GGCCTGCTGGGAACCCGCTGGGCTCGAGAGAGCAGATCTCCACTGTGAGCATG 199  
 OY 259 GATGAGATCTTCCGATCATCATGATGAGAGAGCAGATCTCCACTGTGAGCATG 318  
 DB 200 GATGAGATCTTCCGATCATCATGATGAGAGAGCAGATCTCCACTGTGAGCATG 259  
 OY 319 CCGCTCAGAGCCACCTCAGAGAGTGTGAGAGATATGTAACCGGATGAAGAGTA 378  
 DB 260 CCGCTCAGAGCCACCTCAGAGAGTGTGAGAGAGTGTGTAACCGGATGAAGAGTA 319  
 OY 379 TGGGGAACAGATTGATCTTCCACCCATCGCAGAGATACGTACGTAGTGGCCGGA 438  
 DB 320 -CGGGAACAGATTGA-CCTTCTACCCACCGCAGAGATACGTAGTGGCCGGA 377  
 OY 439 TGGGAATCCAGTGAATGAGATGAGAGATCGAATGAGATCAATGAGTGGCCGAC 498

```

Db 378 TGGAGATCCAGTGAAGTACTTACAGGCGATCCGAATTGACTAGATATCANGCGCAC 437
OY 499 CCTCAAGTTGGCTGTGAGAGCATTTGGAGAGATAGAGATGAATGAATTCATT 558
    |||||
Db 438 CCTCAAGTTGGCTGTGAGAGCATTTGGAGAGATAGAGATGAATGAATTCATT 497
OY 559 TTCCGAGAGGCTGACATGTTAAAGACAACCTTTCAGTAAAGCGACAGATCTTGTGA 618
    |||
Db 498 CTCAGAGAGGCTGACACAGTTAAAGACAACCTTTCAGTAAAGCGACAGATCTATGTGA 557
OY 619 CCATGCCCTGCACATATC 636
    |||
Db 558 CCATGCCCTGCACAGATC 575

```

Search completed: March 30, 2003, 18:47:31  
 Job time : 197.291 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:39:57 ; Search time 44.7061 Seconds  
(without alignments)  
5529.028 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806  
Sequence: 1 cggcccaaggctggggccaa.....aagagagagagagaacta 806

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/lna/5A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/lna/5B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/lna/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	48	6.0	1864	4	US-09-149-476-130 Sequence 130, App1
2	44.6	5.5	1395	2	US-08-553-367A-1 Sequence 1, App1
3	44.6	5.5	1395	4	US-09-295-306-1 Sequence 1, App1
4	44.6	5.5	1395	4	US-09-734-719-1 Sequence 1, App1
5	44.6	5.5	7218	1	US-08-232-463-14 Sequence 14, App1
6	43.8	5.4	1813	5	PCT-US94-12883-3 Sequence 3, App1
7	43.8	5.4	1898	1	US-08-342-411A-1 Sequence 1, App1
8	43.2	5.4	2296	4	US-08-496-841C-137 Sequence 137, App1
9	42.2	5.2	3214	1	US-08-484-105-17 Sequence 17, App1
10	42.2	5.2	3214	1	US-08-484-105-17 Sequence 17, App1
11	42.2	5.2	3300	4	US-09-336-643A-82 Sequence 82, App1
12	41.6	5.2	3300	4	US-09-385-982-389 Sequence 389, App1
13	41.6	5.2	2246	4	US-09-363-708-3 Sequence 3, App1
14	41.6	5.2	7859	1	US-07-854-596B-4 Sequence 4, App1
15	41.6	5.2	7859	2	US-08-450-905B-15 Sequence 15, App1
16	41.6	5.2	7859	3	US-07-982-759F-15 Sequence 15, App1
17	41.4	5.1	1075	4	US-08-400-006B-6 Sequence 6, App1
18	41.4	5.1	1212	4	US-09-182-145-34 Sequence 34, App1
19	41.4	5.1	1212	4	US-09-182-145-35 Sequence 35, App1
20	41.4	5.1	1582	3	US-08-545-196B-10 Sequence 10, App1
21	41.4	5.1	1582	3	US-08-545-196B-12 Sequence 12, App1
22	41.4	5.1	1193	4	US-09-372-422A-23 Sequence 23, App1
23	41.4	5.1	1737	1	US-08-202-056-4 Sequence 4, App1
24	41.4	5.1	1737	1	US-08-076-093A-3 Sequence 3, App1
25	41.4	5.1	1737	1	US-08-701-265-3 Sequence 3, App1
26	41.4	5.1	1737	1	US-08-284-586-3 Sequence 3, App1
27	41.4	5.1	1737	2	US-08-805-478-3 Sequence 3, App1

28	41.4	5.1	1737	2	US-08-802-627A-3 Sequence 3, App1
29	41.4	5.1	1737	2	US-08-801-238-3 Sequence 3, App1
30	41.4	5.1	1737	2	US-08-801-228-3 Sequence 3, App1
31	41.4	5.1	1737	3	US-09-104-296-3 Sequence 3, App1
32	41.4	5.1	1737	5	PCT-US94-06380-2 Sequence 2, App1
33	40.6	5.0	1117	4	US-09-247-373B-33 Sequence 33, App1
34	40.6	5.0	1639	2	US-08-737-524B-1 Sequence 1, App1
35	40.6	5.0	2186	4	US-09-360-545-66 Sequence 66, App1
36	40.4	5.0	1129	4	US-09-227-357-40 Sequence 40, App1
37	40.2	5.0	140	1	US-08-628-417-5 Sequence 5, App1
38	40.2	5.0	240	1	US-08-628-417-6 Sequence 6, App1
39	40.2	5.0	1485	4	US-09-372-422A-39 Sequence 39, App1
40	40.2	5.0	2205	3	US-08-888-077A-41 Sequence 41, App1
41	40.2	5.0	2301	4	US-09-332-191-8 Sequence 8, App1
42	40.2	5.0	2301	4	US-09-232-200-8 Sequence 8, App1
43	40.2	5.0	2301	4	US-09-232-197-8 Sequence 8, App1
44	40.2	5.0	2301	4	US-09-232-201-8 Sequence 8, App1
45	40.2	5.0	2710	4	US-09-232-200-44 Sequence 44, App1

## ALIGNMENTS

RESULT 1  
US-09-149-476-130  
Sequence 130, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23





EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 6.0%; Score 48; DB 4; Length 1864;  
Best Local Similarity 63.2%; Pred. No. 0.0012;  
Matches 72; Conservative 1; Mismatches 41; Indels 0; Gaps 0;

OY 690 CCAGAGGGGAAATGTCGCAATGCTTTATATATATGTTTACTGAAATTAAT 749  
DB 1747 CCAGGGGTCGAAATTTTATATATGTCATGATTAAGTTTGTGAAAKAAA 1806  
OY 750 GAAATATGAAACCAAGTAAACCAAGAGAGAGAGAGAGA 803  
DB 1807 AA 1860

## RESULT 2

US-08-553-367A-1  
Sequence 1, Application US/08553367A

PATENT No. 5939539

GENERAL INFORMATION:  
APPLICANT: Theodor LANGE et al.  
TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)

NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
STREET: 2033 K Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 mb

COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/553,367A  
FILING DATE: No. 5939539ember 27, 1995

CLASSIFICATION: 800  
PRIOR APPLICATION DATA:

APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Lee Cheng

REGISTRATION NUMBER: 40,949  
REFERENCE/DOCKET NUMBER: 49/FD4, 5M2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-721-8200  
TELEFAX: 202-721-8250

TELEX:  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1395 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO  
ANTI-SENSE: NO

ORIGINAL SOURCE:  
ORGANISM: Cucurbita maxima

IMMEDIATE SOURCE:  
CLONE: Clone pB11

US-08-553-367A-1

Query Match 5.5%; Score 44.6; DB 2; Length 1395;  
Best Local Similarity 63.6%; Pred. No. 0.0091;  
Matches 68; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 692 CAGAGGGGAAATGTCGCAATGCTTTATATATGTTTACTGAAATTAATCGA 751  
DB 1289 CTGGAAGTTCGCAATTAATAAACCTTTATATACATGTTCTTACTTGCCTCAAAAA 1348  
OY 752 AAAAAATGAAACCAAGTAAACCAAGAGAGAGAGA 798  
DB 1349 AA 1395

## RESULT 3

US-09-295-306-1  
Sequence 1, Application US/09295306

PATENT No. 6198021

GENERAL INFORMATION:  
APPLICANT: Theodor LANGE et al.

TITLE OF INVENTION: GA 20-OXIDASE GENE SEQUENCES (As Amended)

NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.  
STREET: 2033 K Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 mb

COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS

SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/295,306  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/553,367  
FILING DATE: No. 6198021ember 27, 1995

ATTORNEY/AGENT INFORMATION:  
NAME: Lee Cheng

REGISTRATION NUMBER: 40,949  
REFERENCE/DOCKET NUMBER: 49/DIV-FD4, 5M2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-721-8200  
TELEFAX: 202-721-8250

TELEX:  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1395 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: NO  
ANTI-SENSE: NO

ORIGINAL SOURCE:  
ORGANISM: Cucurbita maxima

IMMEDIATE SOURCE:  
CLONE: Clone pB11

US-09-295-306-1

Query Match 5.5%; Score 44.6; DB 4; Length 1395;  
Best Local Similarity 63.6%; Pred. No. 0.0091;  
Matches 68; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY 692 CAGAGGGGAAATGTCGCAATGCTTTATATATGTTTACTGAAATTAATCGA 751  
DB 1289 CTGGAAGTTCGCAATTAATAAACCTTTATATACATGTTCTTACTTGCCTCAAAAA 1348  
OY 752 AAAAAATGAAACCAAGTAAACCAAGAGAGAGAGA 798  
DB 1349 AA 1395

```

      RESULT 4
US-09-734-719-1
; Sequence 1, Application US/09734719
; Patent No. 645675
; GENERAL INFORMATION:
; APPLICANT: Theodor LANGE et al.
; TITLE OF INVENTION: GA-20-OXIDASE GENE SEQUENCES (As Amended)
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/734,719
; FILING DATE: 13-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/285,306
; FILING DATE: April 21, 1999
; APPLICATION NUMBER: 08/553,367
; FILING DATE: NO. 645675ember 27, 1995
; APPLICATION NUMBER: PCT/EP94/01664
; FILING DATE: May 24, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 2000-1678/LC/01784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cucurbita maxima
; IMMEDIATE SOURCE:
; CLONE: Clone PB11
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-734-719-1

Query Match          5.5%; Score 44.6; DB 4; Length 1395;
Best Local Similarity 63.6%; Pred.No. 0.0091;
Matches 68; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY   692 CAGGAGGGCAAAATGCTGCATATGCCCTTTATATATATATGTATTACGTAATAATTACTGA 751
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db    1289 CTGGAAGTGTGCATTAATAATTAACCCCTTAAATCATCAAGTCCTTACTTGCTCAAAAAA 1348

OY   752 AAAAATATGAAACCAAAGTAAAAAAAAAAAAAAAAAAGAAGAGAGAGA 798
      |||||  ||  ||||  |||||||||||||||||||
Db    1349 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1395

RESULT 5
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
```

[illegible]



```

1 FILING DATE: 28-Jun-1995
2 CLASSIFICATION: <Unknown>
3 ATTORNEY/AGENT INFORMATION:
4 NAME: Paul F. Fehner, Ph.D.
5 REGISTRATION NUMBER: 35,135
6 TELECOMMUNICATION INFORMATION:
7 TELEPHONE: (212) 527-7700
8 TELEFAX: (212) 753-6237
9 INFORMATION FOR SEQ ID NO: 137:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 2296 base pairs
12 TYPE: nucleic acid
13 STRANDEDNESS: single
14 TOPOLOGY: linear
15 MOLECULE TYPE: CDNA
16 SEQUENCE DESCRIPTION: SEQ ID NO: 137:
17 US-08-496-841C-137
18
19 Query Match 5.4%; Score 43.2; DB 4; Length 2296;
20 Match Local Similarity 48.8%; Pred. No. 0.027;
21 Matches 63; Conservative 8; Mismatches 58; Indels 0; Gaps 0.
22
23 Oy 677 TTGAAGATTCACCCCGAGAGGGGAAATATGCGCAATGCTTTATATATATGTTT 736
24 Db 2167 KTTWCGGAARACACACNNCGGCGNGNGNNTTAAATATTAATCCTAAATATATNTANT 2226-
25 Oy 737 ACTGAATTAATCTGAAAAATATGTAACCAAAAGTAAAAATAGAGAGAGAGA 796
26 Db 2227 CCTKTTAANAAAAATATATATATATATATATATATATATATATATATATATAT 2286
27 Oy 797 GAGAGACT 805
28 Db 2287 AARGAATT 2295
29
30 RESULT 9
31 US-08-484-105-17
32 Sequence 17 Application US/08484105
33 Patent No. 5589341
34 GENERAL INFORMATION:
35 APPLICANT: STILLMAN, Bruce
36 APPLICANT: BELL, Stephen P
37 APPLICANT: KOBAYASHI, Ryuji
38 APPLICANT: RINE, Jasper
39 APPLICANT: FOSS, Margit
40 APPLICANT: McNALLY, Francis J
41 APPLICANT: LAURENSEN, Patricia
42 APPLICANT: HERSKOWITZ, Ira
43 APPLICANT: LI, Joachim J
44 APPLICANT: GAVIN, Kimberly
45 TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
46 NUMBER OF SEQUENCES: 24
47 CORRESPONDENCE ADDRESSES:
48 ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
49 STREET: 4 Embarcadero Center, Suite 3400
50 CITY: San Francisco
51 STATE: California
52 COUNTRY: USA
53 ZIP: 94111-4187
54
55 COMPUTER READABLE FORM:
56 MEDIUM TYPE: Floppy disk
57 COMPUTER: IBM PC compatible
58 OPERATING SYSTEM: PC-DOS/MS-DOS
59 SOFTWARE: PatentIn Release #1.0, Version #1.30
60 CURRENT APPLICATION DATA:
61 APPLICATION NUMBER: US/08/484,105
62 FILING DATE:
63 CLASSIFICATION: 435
64 ATTORNEY/AGENT INFORMATION:
65 NAME: Osman Ph.D., Richard Aron
66 REGISTRATION NUMBER: 36,627
67 REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
68 TELECOMMUNICATION INFORMATION:

```

```

: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3214 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 220..2802
: US-08-484-105-17

Query Match 5.2%; Score 42.2; DB 1; Length 3214;
Best Local Similarity 65.3%; Pred. No. 0.06;
Matches 62; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Oy 704 ATGGGCGCAATCCCTTTATATATATATGTTTACTGCAATTAAGTGAAGAAATATGAAA 763
Db 3119 ATTTTGGAATTTTCCTTAATTAATTTGCTTTTAAAAAATTTTTTTTTTTTTTTT 3178

Oy 764 CCAAGAGTAAATATATATATATATATATATATATATATATATATATATATATATAT 798
Db 3179 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3213

RESULT 10
US-08-484-106-17
: Sequence 17, Application US/08484106
: Patent No. 5614618
: GENERAL INFORMATION:
: APPLICANT: STILLMAN, Bruce
: APPLICANT: BELL, Stephen P
: APPLICANT: KOBAYASHI, Ryuji
: APPLICANT: RINE, Jasper
: APPLICANT: FOSS, Margit
: APPLICANT: MCNALLY, Francis J
: APPLICANT: LAURENSEN, Patricia
: APPLICANT: HERSKOWITZ, Ira
: APPLICANT: LI, Joachim J
: APPLICANT: GAVIN, Kimberly
: TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,106
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman Ph.D., Richard Aron
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-59032/DJB/BAO
: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3214 base pairs
: TYPE: nucleic acid

```

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 220..2802  
US-08-484-106-17

Query Match  
Best Local Similarity 5.2%; Score 42.2; DB 1; Length 3214;  
Matches 62; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 704 ATGCTGCAATGCTTTATATATATGTTTACTGAAATTAAGTGAATATGAA 763  
Db 3119 ATTTGGAATTTTCATTAAATATGCTTTTAAATATATATATATATATAT 3178  
QY 764 CCAAAAGTAAAAAGAGAGAGAGAGA 798  
Db 3179 AAAAAAAAAAAAAAAAAAAAAAAAAA 3213

## RESULT 11

US-09-336-643A-82  
Sequence 82, Application US/09336643A  
General No. 6399761  
GENERAL INFORMATION:  
APPLICANT: Miller, Andrew P.  
APPLICANT: Curran, Mark Edward  
APPLICANT: Hu, Ping  
APPLICANT: Rutger, Marc  
APPLICANT: Wang, Jian-Wang  
TITLE OF INVENTION: NO. 6399761e1 Human Potassium Channels  
FILE REFERENCE: SEQ-15P  
CURRENT APPLICATION NUMBER: US/09/336,643A  
PRIOR APPLICATION NUMBER: 60/076,687  
PRIOR FILING DATE: 1998-06-07  
PRIOR APPLICATION NUMBER: 60/116,448  
PRIOR FILING DATE: 1999-01-19  
PRIOR APPLICATION NUMBER: PCT/US99/03826  
PRIOR FILING DATE: 1999-02-22  
NUMBER OF SEQ ID NOS: 87  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 82  
LENGTH: 3300  
TYPE: DNA  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (50)...(1285)  
US-09-336-643A-82

Query Match  
Best Local Similarity 5.2%; Score 42.2; DB 4; Length 3300;  
Matches 62; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 708 TGGCAATGCTTTATATATATGTTTACTGAAATTAAGTGAATATGAA 767  
Db 3206 TGTGAATGTTGACAAATTAAGTTGTTAATTAAGTCAAAAAAAAAAAAA 3265  
QY 768 AAGTAAAAAGAGAGAGAGAGAGA 802  
Db 3266 AAAAAAAAAAAAAAAAAAAAAAAAAA 3300

## RESULT 12

US-09-385-982-389/c  
Sequence 389, Application US/09385982  
General No. 6262334  
GENERAL INFORMATION:  
APPLICANT: ENDEGE, WILSON O., ET AL.  
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
TITLE OF INVENTION: PRODUCTS: 11

FILE REFERENCE: CCDNA-260XX  
CURRENT APPLICATION NUMBER: US/09/385,982  
CURRENT FILING DATE: 1999-08-30  
EARLIER APPLICATION NUMBER: 09/328,111  
EARLIER FILING DATE: 1999-06-08  
EARLIER APPLICATION NUMBER: 60/117,393  
EARLIER FILING DATE: 1999-01-27  
EARLIER APPLICATION NUMBER: 60/098,639  
EARLIER FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 544  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 389  
LENGTH: 629  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(629)  
OTHER INFORMATION: n = A,T,C or G  
US-09-385-982-389

Query Match  
Best Local Similarity 5.2%; Score 41.6; DB 4; Length 629;  
Matches 68; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 689 CCCAGAGAGGGAATGTCGCAATGCTTTATATATATGTTTACTGAAATTAAC 748  
Db 119 CACCATGATGTGAATGTTTCACAAATCTTCAAAATATATATGTTTGAACG 60  
QY 749 TGAATAATATGAATCAAAAGTAAAAAGAGAGAGAGAGAGA 803  
Db 59 CAAAAAAAAAAAAAAAAAAAAAAAAA 5

## RESULT 13

US-09-363-708-3  
Sequence 3, Application US/09363708  
General No. 6399747  
GENERAL INFORMATION:  
APPLICANT: Schmidt, et al.  
TITLE OF INVENTION: NOVEL SHC BINDING PROTEIN.  
NUMBER OF SEQUENCES: 12  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/363,708  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 01017/34451  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2246 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
DESCRIPTION: /desc = "mouse PAL CDNA"

US-09-363-708-3

Query Match 5.2%; Score 41.6; DB 4; Length 2246;  
Best Local Similarity 64.6%; Pred. No. 0.074;  
Matches 62; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 708 TGGCATGCGCTTATATATATGTTTACGAAATTAACGAAAAATGAAACCAA 767  
DB 2104 TGGCTGTGATATATATGATGACTGACCAAAAAAAAAAAAAAAAAAAAAA 2163  
OY 768 AAGTAAAAAAAAAAAAAAAAAGAGAGAGAGAGAA 803  
DB 2164 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2199

RESULT 14

US-07-854-596B-4  
; Sequence 4, Application US/07854596B  
; Patent No. 5434073  
; GENERAL INFORMATION:  
; APPLICANT: Dawson, Keith M  
; APPLICANT: Hunter, Michael G  
; APPLICANT: Caplewski, Lloyd G  
; TITLE OF INVENTION: Proteins and nucleic acids  
; NUMBER OF SEQUENCES: 73  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dr. John J. McDonnell  
; STREET: Ten South Wacker Drive, Suite 3000  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/854, 596B  
; FILING DATE: 03-JUN-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McDonnell, John J  
; REGISTRATION NUMBER: 26, 949  
; REFERENCE/DOCKET NUMBER: 92,337  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; TELEX: 910-221-5317  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7859 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..7859  
; OTHER INFORMATION: /note= "sequence of plasmid pSWC"  
US-07-854-596B-4

Query Match 5.2%; Score 41.6; DB 1; Length 7859;  
Best Local Similarity 70.0%; Pred. No. 0.13;  
Matches 56; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY 713 ATGCTTTATATATATCTTTTCTGAAATTAAGTGAACCAAGTAA 772  
DB 5584 ATTCCTTTTATGATATTTGTACAAAAAAAAAAAAAAAAAAAAA 5643  
OY 773 AAAAAAAAAAAAAAAAAAGAGA 792  
DB 5644 AAAAAAAAAAAAAAAAAAAAAA 5663

RESULT 15  
US-08-450-905B-15  
; Sequence 15, Application US/08450905B  
; Patent No. 5856301  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Stem Cell Inhibiting Proteins  
; NUMBER OF SEQUENCES: 178  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HALE and DORR  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; ZIP: 02109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,905B  
; FILING DATE: 26-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/982,759  
; FILING DATE: 08-MAR-1993  
; PRIOR APPLICATION NUMBER: GB 9127319.3  
; FILING DATE: 23-DEC-1991  
; APPLICATION DATA:  
; APPLICATION NUMBER: GB 9221587.0  
; FILING DATE: 14-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BAKER, HOLIE L.  
; REGISTRATION NUMBER: 31,321  
; REFERENCE/DOCKET NUMBER: 102.378.120DV-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-526-6110  
; TELEFAX: 617-526-5000  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7859 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA  
US-08-450-905B-15

Query Match 5.2%; Score 41.6; DB 2; Length 7859;  
Best Local Similarity 70.0%; Pred. No. 0.13;  
Matches 56; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY 713 ATGCTTTATATATATCTTTTCTGAAATTAAGTGAACCAAGTAA 772  
DB 5584 ATTCCTTTTATGATATTTGTACAAAAAAAAAAAAAAAAAAAAA 5643  
OY 773 AAAAAAAAAAAAAAAAAAGAGA 792  
DB 5644 AAAAAAAAAAAAAAAAAAAAAA 5663

Search completed: March 30, 2003, 20:59:05  
Job time : 76.7061 secs





```

DB 172 CTTTCGCGGGGCTCTGCTGGAAACCCCTGCGTCCGAGAGACCCAGATCTCCACCTGTGA 231
OY 191 GCATCGAGGGCTCTGCTGGATGATGAAATGGAAATTTGCCAGGTGAGCCCAAGAAG 250
DB 232 GCATCGAGGGCTCTGCTGGATGATGAAATGGAAATTTGCCAGGTGAGCCCAAGAAG 291
OY 251 ACCATTGAGATGGGATCTTTCCGGATTCATTCAGATGGAGCCAGTCAGTGTGAGGTG 310
DB 292 ACCATTGAGATGGGATCTTTCCGGATTCATTCAGATGGAGCCAGTCAGTGTGAGGTG 351
OY 311 CCTTATGCGCGCTCAGAGAGCCACCTCAGAGAGTGTGAGAGATATGATCCGGATG 370
DB 352 CCTTATGCGCGCTCAGAGAGCCACCTCAGAGAGTGTGAGAGATATGATCCGGATG 411
OY 371 AAGAGATATGGGAAACAGATGATCTTTCCACCCATTCGCAAGACTACGTAGTGTAGTG 430
DB 412 AAGGAGTATGGGAAACAGATGATCTTTCCACCCATTCGCAAGACTACGTAGTGTAGTG 471
OY 431 GCGCGGATGAGAGATCCAGTGAATGAGACCTTGTGAGAGATGAGAGATGAGTCAATT 490
DB 472 GCGCGGATGAGAGATCCAGTGAATGAGACCTTGTGAGAGATGAGAGATGAGTCAATT 531
OY 491 AGCGGACCCCTCAAGTTCGCTGTGAGAGATGAGAGATGAGAGATGAGTCAATT 550
DB 532 AGCGGACCCCTCAAGTTCGCTGTGAGAGATGAGAGATGAGAGATGAGTCAATT 591
OY 551 GAATTCCTTTCCGAGAGAGCTGACATGTTAAAGCAAACTTTGAGATGAGGAAAGAT 610
DB 592 GAATTCCTTTCCGAGAGAGCTGACATGTTAAAGCAAACTTTGAGATGAGGAAAGAT 651
OY 611 CTTTGTGACATGAGCCCTGACATATCGCATGATGAGCTATGAGACCTGAGAGACCCAC 670
DB 652 CTTTGTGACATGAGCCCTGACATATCGCATGATGAGCTATGAGACCTGAGAGACCCAC 711
OY 671 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
DB 712 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
OY 731 GTTTTTACTGAAATTAAGTGAATAATGAAACCAAAAGTA 772
DB 772 GTTTTTACTGAAATTAAGTGAATAATGAAACCAAAAGTA 813

```

```

PRIORITY FILING DATE: 1997-08-26
PRIORITY APPLICATION NUMBER: 60/059113
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059115
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059117
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059122
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059184
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/059263
PRIORITY FILING DATE: 1997-09-18
PRIORITY APPLICATION NUMBER: 60/059352
PRIORITY FILING DATE: 1997-09-19
PRIORITY APPLICATION NUMBER: 60/059588
PRIORITY FILING DATE: 1997-09-19
PRIORITY APPLICATION NUMBER: 60/059836
PRIORITY FILING DATE: 1997-09-24
PRIORITY APPLICATION NUMBER: 60/062250
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/062285
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/062287
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/062814
PRIORITY FILING DATE: 1997-10-24
PRIORITY APPLICATION NUMBER: 60/062816
PRIORITY FILING DATE: 1997-10-24
PRIORITY APPLICATION NUMBER: 60/063045
PRIORITY FILING DATE: 1997-10-24
PRIORITY APPLICATION NUMBER: 60/063082
PRIORITY FILING DATE: 1997-10-31
PRIORITY APPLICATION NUMBER: 60/063127
PRIORITY FILING DATE: 1997-10-24
PRIORITY APPLICATION NUMBER: 60/063327
PRIORITY FILING DATE: 1997-10-27
PRIORITY APPLICATION NUMBER: 60/063329
PRIORITY FILING DATE: 1997-10-27
PRIORITY APPLICATION NUMBER: 60/063550
PRIORITY FILING DATE: 1997-10-28
PRIORITY APPLICATION NUMBER: 60/063561
PRIORITY FILING DATE: 1997-10-28
PRIORITY APPLICATION NUMBER: 60/063704
PRIORITY FILING DATE: 1997-10-29
PRIORITY APPLICATION NUMBER: 60/063733
PRIORITY FILING DATE: 1997-10-29
PRIORITY APPLICATION NUMBER: 60/063735
PRIORITY FILING DATE: 1997-10-29
PRIORITY APPLICATION NUMBER: 60/063738
PRIORITY FILING DATE: 1997-10-29
PRIORITY APPLICATION NUMBER: 60/063755
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/064248
PRIORITY FILING DATE: 1997-11-03
PRIORITY APPLICATION NUMBER: 60/064809
PRIORITY FILING DATE: 1997-11-07
PRIORITY APPLICATION NUMBER: 60/065186
PRIORITY FILING DATE: 1997-11-12
PRIORITY APPLICATION NUMBER: 60/065846
PRIORITY FILING DATE: 1997-11-17
PRIORITY APPLICATION NUMBER: 60/066364
PRIORITY FILING DATE: 1997-11-21
PRIORITY APPLICATION NUMBER: 60/066453
PRIORITY FILING DATE: 1997-11-24
PRIORITY APPLICATION NUMBER: 60/066511
PRIORITY FILING DATE: 1997-11-24
PRIORITY APPLICATION NUMBER: 60/066770
PRIORITY FILING DATE: 1997-11-24
PRIORITY APPLICATION NUMBER: 60/069212
PRIORITY FILING DATE: 1997-12-11
PRIORITY APPLICATION NUMBER: 60/069278
PRIORITY FILING DATE: 1997-12-11

```

PRIOR APPLICATION NUMBER: 60/069334  
 PRIOR FILING DATE: 1997-12-11  
 PRIOR APPLICATION NUMBER: 60/069694  
 PRIOR FILING DATE: 1997-12-16  
 PRIOR APPLICATION NUMBER: 60/072320  
 PRIOR FILING DATE: 1998-01-23  
 PRIOR APPLICATION NUMBER: 60/073612  
 PRIOR FILING DATE: 1998-02-04  
 PRIOR APPLICATION NUMBER: 60/074086  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/074092  
 PRIOR FILING DATE: 1998-02-09  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-02-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081695  
 PRIOR FILING DATE: 1998-04-14  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081818  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082999  
 PRIOR FILING DATE: 1998-04-24  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085149  
 PRIOR FILING DATE: 1998-05-12  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/086414  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/086430  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/087106  
 PRIOR FILING DATE: 1998-05-28  
 PRIOR APPLICATION NUMBER: 60/088026  
 PRIOR FILING DATE: 1998-06-04  
 PRIOR APPLICATION NUMBER: 60/088730  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088741  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090538  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
 Best Local Similarity 98.9%; Pred. No. 5,6e-200;  
 Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGCGCAAGTGAAGTCCAGCGGTCTGCCAGCCCTTGGCCAGCGGGGGGCGCTGG 70  
 450 CTGGGCGCAAGTGAAGTCCAGCGGTCTGCCAGCCCTTGGCCAGCGGGGGGCGCTGG 509  
 71 GACCAAGGTGAGCAACCCGTTTCCCTAATGAAGAAAGCTGGAGTGGCTGGCCCTG 130  
 510 GACCAAGGTGAGCAACCCGTTTCCCTAATGAAGAAAGCTGGAGTGGCTGGCCCTG 569  
 131 CTTCCTGGGCGCTGCTGGGAACCGCTGGCTGGAGAGAGCCAGATCTCCAGTGGGA 190  
 570 CTTCCTGGGCGCTGCTGGGAACCGCTGGCTGGAGAGAGCCAGATCTCCAGTGGGA 629  
 191 GCATGAGGCGCTGCTGGATGAATGAAGTGGAAATTTGCCAGGTGGACCCCAAGAG 250  
 630 GCATGAGGCGCTGCTGGATGAATGAAGTGGAAATTTGCCAGGTGGACCCCAAGAG 689  
 251 ACCATTGAGTGGATCTTCCGATCAATCCAGTGGACCCAGTCAATGATGGAGGTG 310  
 690 ACCATTGAGTGGATCTTCCGATCAATCCAGTGGACCCAGTCAATGATGGAGGTG 749  
 311 CCTTATGCGCGCTGAGAGCCGACCTTCACAGAGCTGCTGAGAGATATGACCGGATG 370  
 750 CCTTATGCGCGCTGAGAGCCGACCTTCACAGAGCTGCTGAGAGATATGACCGGATG 809  
 371 AAGAGTATGGGGAACAGATTGATCTTCCACCCATGCAAGAACTGACGATGAGTG 430  
 810 AAGAGTATGGGGAACAGATTGATCTTCCACCCATGCAAGAACTGACGATGAGTG 869  
 431 GCGCGAATGAGAAATCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 490  
 870 GCGCGAATGAGAAATCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 929  
 491 AGCGGACCCCTCAAGTTTGGTGGAGAGATGGAGAAATGAGAGATGAATCTCATT 550  
 930 AGCGGACCCCTCAAGTTTGGTGGAGAGATGGAGAAATGAGAGATGAATCTCATT 989  
 551 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTCAGTAAAGCAAGAT 610  
 990 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTCAGTAAAGCAAGAT 1049  
 611 CTTTGTGACCATGCTGACATATGCAATGATGAGCTATGAACCACTGGAGACCCAC 670

```

Db      1050  CTTTGTGACCAATGCCCTGCACATATCGCATGTAGAGCTATGAACCACTGGACAGCCAC 1109
|
|
|
Oy      671   ACTGCGTATGATGCATACACCCACGAGAGGGGAAATATGCTGGCAATGCTTTTATATTTAT 730
|
|
|
Db      1110  ACTGCGCTGATGATGCATACCCCGACGAGGGGGAATATGCTGGCAATGCTTTTATATTTAT 1169
|
|
|
Oy      731   GTTTTACTGAATTTAACTGAATAAATATGAACCAAAAGT 771
|
|
|
Db      1170  GTTTTACTGAATTTAACTGAATAAATATGAACCAAAAGT 1210
|
|
|
RESULT 3
US-10-121-049-173
; Sequence 173, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Mei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zhenli
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-173

Query Match          93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5,6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy      11   CTGGGGCCCAAGTGAATCCAGCGGCTGCGCAGGCGTTGGGCCAAGCGCGGCGCCCTGG 70
Db      450   CTGGGGCCCAAGTGAATCCAGCGGCTGCGCAGGCGTTGGGCCAAGCGCGGCGCCCTGG 509
|
|
|
Oy      71   GACCAAGGTGAGAGCAACCCCGTTACCCCTAAARATGAAGGCTGGGGTTGCTGACCCCTG 130
|
|
|
Db      510   GAGCAGAGGTGAGAGGACCCCATTAACGCTAAAGATGAAGAGGTGGGGTTGGCTGGCCCTG 569
|
|
|
Oy      131   CTTCTGGGGGCGCCCTGCTGGGAACCGCCCTGGGCTCGGAGAGGACGACGATCTCCATGTGGA 190
Db      570   CTTCTGGGGGCGCCCTGCTGGGAACCGCCCTGGGCTCGGAGAGGACGACGATCTCCATGTGGA 629
|
|
|
Oy      191   GCATTCAGAGGCGCTGCTGGTGAATGAATGAATGGGAAATTTGCCAGGTGGAGACCCCAAGAG 250
Db      630   GCATTCAGAGGCGCTGCTGGTGAATGAATGAATGGGAAATTTGCCAGGTGGAGACCCCAAGAG 689
|
|
|
Oy      251   ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGCTCAGTGTGGAGGTG 310
Db      690   ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGCTCAGTGTGGAGGTG 749
|
|
|
Oy      311   CCTTATGCCCGCTCAGAGAGCCCACTCAGACAGCTGCTGGAGAGATATGTACCGGATG 370
|
|
|

```

DB	750	CCATTAGCCCGCTCAGAGGCCCCACCTCACAGAGCTGCTGGAGGAGATATGTGACCCGGATG	809
QY	371	AAGGAGTATGGGGAACAGATTTATCTTCCACCCATCGCAAGAACTACTAGTGTAGTG	430
DB	810	AAGGATATAGGGGAACAGATTTATCTTCCACCCATCGCAAGAACTACTAGTGTAGTG	869
QY	431	GGCCGGAATGAGAAATCCAGTGAACCTGACCTCAAGGCAATCCGATTCAGTACATAT	490
DB	870	GGCCGGAATGAGAAATCCAGTGAACCTGACCTCAAGGCAATCCGATTCAGTACATAT	929
QY	491	ACGGCACCCTCAAGTTTCCGTGTGAGACATTTGGAGGAATAGAGATGAATCTATT	550
DB	930	ACGGCACCCTCAAGTTTCCGTGTGAGACATTTGGAGGAATAGAGATGAATCTATT	989
QY	551	GAATTCCTTTCCCGAGAGGCTGCATATGTTAAAGACAACTTTGGAGTAAAGCAAT	610
DB	990	GAATTCCTTTCCCGAGAGGCTGCATATGTTAAAGACAACTTTGGAGTAAAGCAAT	1049
QY	611	CTTTGTGACCATGCCCTGCACATATTCGATGATGACCTATGAACCACTGGAGCGCCAC	670
DB	1050	CTTTGTGACCATGCCCTGCACATATTCGATGATGACCTATGAACCACTGGAGCGCCAC	1109
QY	671	ACTGGCTTATGATCACCCTCCAGAGGGGAAATGCTGGCAATGCTTTTATATTTAT	730
DB	1110	ACTGGCTTATGATCACCCTCCAGAGGGGAAATGCTGGCAATGCTTTTATATTTAT	1169
QY	731	CTTTTACTGAATTAACCTGAAAAAATATGAACCAAAAGT	771
DB	1170	GTTTTTACTGAATTAACCTGAAAAAATATGAACCAAAAGT	1210
RESULT 4			
US-10-123-904-173			
; Sequence 173, Application US/10123904			
; Publication No. US20030022328A1			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin P.			
; APPLICANT: Beresini, Maureen			
; APPLICANT: DeForge, Laura			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Gao, Wei-Qiang			
; APPLICANT: Gerritsen, Mary E.			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Gurney, Austin L.			
; APPLICANT: Sherwood, Steven			
; APPLICANT: Smith, Victoria			
; APPLICANT: Stewart, Timothy A.			
; APPLICANT: Tumas, Daniel			
; APPLICANT: Watanabe, Colin K			
; APPLICANT: Zhang, Zhen			
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
; FILE OF INVENTION: ACIDS ENCODING THE SAME			
; FILE REFERENCE: P3330R1C54			
; CURRENT APPLICATION NUMBER: 2002-04-16			
; PRIORITY FILING DATE: 2002-04-16			
; Prior Application removed - See file Wrapper or Palm			
; NUMBER OF SEQ ID NOS: 550			
; SEQ ID NO 173			
; LENGTH: 1210			
; TYPE: DNA			
; ORGANISM: Homo Sapien			
US-10-123-904-173			
Query Match			
Best Local Similarity 93.0%; Score 749.4; DB 9; Length 1210;			
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;			
QY	11	CTGGGGCCAAAGTGAATCCACGCGGTCTGCGCAGGCGTTGGCCACAGCGCGCCCTGG	70
DB	450	CTGGGGCCAAAGTGAATCCACGCGGTCTGCGCAGGCGTTGGCCACAGCGCGCGCCCTGG	509

Qy	71	GACC	AAAGGGGAGCAACCC	CCCTTACCTTAAARATGA	AAAGCTGGGGTGGCCCTG	130
Db	510	GAC	ACAGGGGGAGCCACCC	CTTACGCTAAAGATGA	AAAGGCTGGGGTGGCCCTG	569
Qy	131	CTTCTG	GGGGCCCTGCTGGG	ACACGCTGGGCTGGAGAC	CCGAGATTCACACTGGGA	190
Db	570	CTTCTG	GGGGCCCTGCTGGG	ACACGCTGGGCTGGAGAC	CCGAGATTCACACTGGGA	629
Qy	191	GCATG	AGGGCTCTGCTGGAT	ATGACTAGATGGGAA	TTGGCCAGGTGACCCCAAG	250
Db	630	GCATG	AGGGCTCTGCTGGAT	ATGACTAGATGGGAA	TTGGCCAGGTGACCCCAAG	689
Qy	251	ACCAT	TCAGATGGATCTTTCCG	ATCAATCCAGATGGG	ACGACAGTACGTGGAGGTG	310
Db	690	ACCAT	TCAGATGGATCTTTCCG	ATCAATCCAGATGGG	ACGACAGTACGTGGAGGTG	749
Qy	311	CTTTATG	CCCGCTCAGAGGCC	CCCTCAGAGCTGCTGG	AGAGATATGTGACC	370
Db	750	CTTTATG	CCCGCTCAGAGGCC	CCCTCAGAGCTGCTGG	AGAGATATGTGACC	809
Qy	371	AAGGAT	ATAGGGGAACAGAT	TGATCTCTTCCACCC	ATGGGAAACTATAGT	430
Db	810	AAGGAT	ATAGGGGAACAGAT	TGATCTCTTCCACCC	ATGGGAAACTATAGT	869
Qy	431	GGCCG	GAATGGGAATCCAG	TGAACTGGACCTAC	AAAGCATCCGAA	490
Db	870	GGCCG	GAATGGGAATCCAG	TGAACTGGACCTAC	AAAGCATCCGAA	929
Qy	491	AGCGC	ACACCTCAAGTTGG	CGTGGAGAGATGTG	AGAAATACAGAGTAC	550
Db	930	AGCGC	ACACCTCAAGTTGG	CGTGGAGAGATGTG	AGAAATACAGAGTAC	989
Qy	551	GAAT	CTTTTCCGAGAGG	CGTGACAAATGTTAA	GACAACTTGGAGTAA	610
Db	990	GAAT	CTTTTCCGAGAGG	CGTGACAAATGTTAA	GACAACTTGGAGTAA	1049
Qy	611	CTTTG	TGACATGCGCCG	ACATATGCGATGAG	ACTATGAAACATGTG	670
Db	1050	CTTTG	TGACATGCGCCG	ACATATGCGATGAG	ACTATGAAACATGTG	1109
Qy	671	ACTG	CTTGATGGATCACC	CCCGAGAGGGGAA	ATGTGGCAATGCTTTT	730
Db	1110	ACTG	CTTGATGGATCACC	CCCGAGAGGGGAA	ATGTGGCAATGCTTTT	1169
Qy	731	GTTTT	ACTGAATTAACATA	AAAAAATATGAA	ACCAAAAGT	771
Db	1170	GTTTT	ACTGAATTAACATA	AAAAAATATGAA	ACCAAAAGT	1210

RESULT 5  
US-10-140-470-173  
: Sequence 173, Application US/10140470  
: Publication No. US20030022331A1  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin P.  
: APPLICANT: Beresini, Maureen  
: APPLICANT: DeForge, Laura  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Gao, Wei-Qiang  
: APPLICANT: Gerritsen, Mary E.  
: APPLICANT: Goddard, Audrey  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Sherwood, Steven  
: APPLICANT: Smith, Victoria  
: APPLICANT: Stewart, Timothy A.  
: APPLICANT: Tumas, Daniel  
: APPLICANT: Watanabe, Colin K  
: APPLICANT: Wood, William  
: APPLICANT: Zhang, Zemin  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-470-173

Query Match      93.0%      Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5.6e-200;
Matches 753; Conservative 1; Mismatches 7; Inde1s 0; Gaps 0

```

RESULT 6  
US-10-175-746-173  
; Sequence 173, Application US/10175746  
; Publication No. US20030027270A1

```

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 173
LENGTH: 1210
TYPE: DNA
ORGANISM: Homo Sapien
US-10-175-746-173

```

```

Query Match          93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5,6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

OY 11 CTGGGGCCAAAGTGAAGATCCAGGCTGCGCAGGCTTGGGCGACGGGGGCGCCCTGG 70
DB 450 CTGGGGCCAAAGTGAAGATCCAGGCTTGGGCGACGGGGGCGCCCTGG 509
OY 71 GACCAAGGTGAGCAACCCCGTTACCTTAARATGAAGGCTGGGCTGGGCTGG 130
DB 510 GAGCAGAGGTGAGGAGGAGCCCATTAAGATGAAGGCTGGGCTGGGCTGG 569
OY 131 CTCTGGGGGCGCTGCTGGGAAACCGCTGCGAGGAGCCAGATTCACACTGTGA 190
DB 570 CTCTGGGGGCGCTGCTGGGAAACCGCTGCGAGGAGCCAGATTCACACTGTGA 629
OY 191 GCATGAGGCGCTGCTGGTGAATAGATGAAGTGGAAATGGCCAGGTGAGCCCAAGAG 250
DB 630 GCATGAGGCGCTGCTGGTGAATAGATGAAGTGGAAATGGCCAGGTGAGCCCAAGAG 689
OY 251 ACCATTGAGATGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCACTGTGAGT 310
DB 690 ACCATTGAGATGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCACTGTGAGT 749
OY 311 CCTTATGCGCGCTCAGAGGCCCACTCAGAGAGCTGCTGGAGAGATATGTAACGGATG 370
DB 750 CCTTATGCGCGCTCAGAGGCCCACTCAGAGAGCTGCTGGAGAGATATGTAACGGATG 809
OY 371 AAGGATATGGGGAACAGATGATCTTCACCATCCGAAAGACTAGTACGTAGT 430
DB 810 AAGGATATGGGGAACAGATGATCTTCACCATCCGAAAGACTAGTACGTAGT 869
OY 431 GCGCGGAATGAGAAATCCAGTGAATGCACTCAAGAGCATCCGATCACTGAT 490
DB 870 GCGCGGAATGAGAAATCCAGTGAATGCACTCAAGAGCATCCGATCACTGAT 929
OY 491 AGCGGACACCTTAAGTTCGCTGTAGAGAGATTTGAGAGAAATACGAGATCACTATT 550
DB 930 AGCGGACACCTTAAGTTCGCTGTAGAGAGATTTGAGAGAAATACGAGATCACTATT 989
OY 551 GAATCTTTTCCGAGAGGCTGACAAATGTTAAGCAAACTTTGAGATGAAGGAACAGAT 610

```

```

DB 990 GAATCTTTTCCGAGAGGCTGACAAATGTTAAGCAAACTTTGAGATGAAGGAGACAGAT 1049
OY 611 CTCTTGACCAATGCCCTGACATATGCAATGATGACATGATGAACCACTGAGAGCCAC 670
DB 1050 CTCTTGACCAATGCCCTGACATATGCAATGATGACATGATGAACCACTGAGAGCCAC 1109
OY 671 ACTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
DB 1110 ACTGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1169
OY 731 GTTTTACGAAATTAAGTGAAGAAATATGAACCAAAAGT 771
DB 1170 GTTTTACGAAATTAAGTGAAGAAATATGAACCAAAAGT 1210

```

RESULT 7  
US-10-176-918-173

```

Sequence 173, Application US/10176918
Publication No. US20030027275A1
GENERAL INFORMATION:

```

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 173
LENGTH: 1210
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-918-173

```

```

Query Match          93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5,6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

OY 11 CTGGGGCCAAAGTGAAGATCCAGGCTGCGCAGGCTTGGGCGACGGGGGCGCCCTGG 70
DB 450 CTGGGGCCAAAGTGAAGATCCAGGCTTGGGCGACGGGGGCGCCCTGG 509
OY 71 GACCAAGGTGAGCAACCCCGTTACCTTAARATGAAGGCTGGGCTGGGCTGG 130
DB 510 GAGCAGAGGTGAGGAGGAGCCCATTAAGATGAAGGCTGGGCTGGGCTGG 569
OY 131 CTCTGGGGGCGCTGCTGGGAAACCGCTGCGAGGAGCCAGATTCACACTGTGA 190
DB 570 CTCTGGGGGCGCTGCTGGGAAACCGCTGCGAGGAGCCAGATTCACACTGTGA 629
OY 191 GCATGAGGCGCTGCTGGTGAATAGATGAAGTGGAAATGGCCAGGTGAGCCCAAGAG 250
DB 630 GCATGAGGCGCTGCTGGTGAATAGATGAAGTGGAAATGGCCAGGTGAGCCCAAGAG 689
OY 251 ACCATTGAGATGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCACTGTGAGT 310
DB 690 ACCATTGAGATGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCACTGTGAGT 749

```

QY 311 CTTATGCCCCGTGAGAGCCCACTCAAGAGCTGTGAGAGATATGACCGAGT 370  
 Db 750 CTTATGCCCCGTGAGAGCCCACTCAAGAGCTGTGAGAGATATGACCGAGT 809  
 QY 371 AAGAGATGAGGAGACATGATGATCTTCCACCACTGCAAGAACATGATGATG 430  
 Db 810 AAGAGATGAGGAGACATGATGATCTTCCACCACTGCAAGAACATGATGATG 869  
 QY 431 GCGCGAATGAGAGATGATGATGATCTTCCACCACTGCAAGAACATGATGATG 490  
 Db 870 GCGCGAATGAGAGATGATGATGATCTTCCACCACTGCAAGAACATGATGATG 929  
 QY 491 AGCGGACCCCTCAAGTGTGCTGTGAGAGATGATGATGATGATGATGATGATG 550  
 Db 930 AGCGGACCCCTCAAGTGTGCTGTGAGAGATGATGATGATGATGATGATGATG 989  
 QY 551 GAATTTCTTTCCGAGAGGCTGACATGTTTAAAGAACATTTTCAATGAGCAAGAT 610  
 Db 990 GAATTTCTTTCCGAGAGGCTGACATGTTTAAAGAACATTTTCAATGAGCAAGAT 1049  
 QY 611 CTTTGTGACATGCGCTGACATGATGATGATGATGATGATGATGATGATGATG 670  
 Db 1050 CTTTGTGACATGCGCTGACATGATGATGATGATGATGATGATGATGATGATG 1109  
 QY 671 ACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 730  
 Db 1110 ACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1169  
 QY 731 GTTTTACTGAAATTAATGAAAAATATGAAACCAAAAGT 771  
 Db 1170 GTTTTACTGAAATTAATGAAAAATATGAAACCAAAAGT 1210

RESULT 8

US-10-176-921-173  
 ; Sequence 173, Application US/10176921  
 ; Publication No. US20030027276A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K  
 ; APPLICANT: Wood, William  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3330RJC288  
 ; CURRENT APPLICATION NUMBER: US/10/176,921  
 ; PRIOR Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 173  
 ; LENGTH: 1210  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-176-921-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
 Best Local Similarity 98.9%; Pred. No. 5,66-200;  
 Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 11 CTGGGCGCAAGATGAAGTCCAGCGGCTGCGAGCGCTTGGGCCACGCGCGCCCTG 70  
 Db 450 CTGGGCGCAAGATGAAGTCCAGCGGCTGCGAGCGCTTGGGCCACGCGCGCCCTG 509  
 QY 71 GACCAAGGATGAGCAACCCCGTTTACCCTTAAATGAGAGAGAGAGAGAGAGAGAG 130  
 Db 510 GAGCAGAGGATGAGCAACCCCGTTTACCCTTAAATGAGAGAGAGAGAGAGAGAGAG 569  
 QY 131 CTTCTGGGGCCCTGCTGGGAGACCGCTGCGCTGGGAGAGAGAGAGAGAGAGAGAG 190  
 Db 570 CTTCTGGGGCCCTGCTGGGAGACCGCTGCGCTGGGAGAGAGAGAGAGAGAGAGAG 629  
 QY 191 GCATGCGGGCTGTGGGATGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250  
 Db 630 GCATGCGGGCTGTGGGATGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689  
 QY 251 ACATTCAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 310  
 Db 690 ACATTCAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 749  
 QY 311 CTTTATGCCCCGTGAGAGCCCACTCAAGAGCTGTGAGAGATATGATGATGATGATG 370  
 Db 750 CTTTATGCCCCGTGAGAGCCCACTCAAGAGCTGTGAGAGATATGATGATGATGATG 809  
 QY 371 AAGAGATGAGGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 430  
 Db 810 AAGAGATGAGGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 869  
 QY 431 GCGCGAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490  
 Db 870 GCGCGAATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 929  
 QY 491 AGCGGACCCCTCAAGTGTGCTGTGAGAGATGATGATGATGATGATGATGATGATG 550  
 Db 930 AGCGGACCCCTCAAGTGTGCTGTGAGAGATGATGATGATGATGATGATGATGATG 989  
 QY 551 GAATTTCTTTCCGAGAGGCTGACATGTTTAAAGAACATTTTCAATGAGCAAGAT 610  
 Db 990 GAATTTCTTTCCGAGAGGCTGACATGTTTAAAGAACATTTTCAATGAGCAAGAT 1049  
 QY 611 CTTTGTGACATGCGCTGACATGATGATGATGATGATGATGATGATGATGATGATG 670  
 Db 1050 CTTTGTGACATGCGCTGACATGATGATGATGATGATGATGATGATGATGATGATG 1109  
 QY 671 ACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 730  
 Db 1110 ACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1169  
 QY 731 GTTTTACTGAAATTAATGAAAAATATGAAACCAAAAGT 771  
 Db 1170 GTTTTACTGAAATTAATGAAAAATATGAAACCAAAAGT 1210

RESULT 9

US-10-137-865-173  
 ; Sequence 173, Application US/10137865  
 ; Publication No. US20030032155A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K



; APPLICANT: Wood,William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P330R1C154  
 ; CURRENT APPLICATION NUMBER: US/10/137,865  
 ; CURRENT FILING DATE: 2002-05-10  
 ; Prior Application removed - See Palm or File Wrapper  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 173  
 ; LENGTH: 1210  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-137-865-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
 Best Local Similarity 98.9%; Pred. No. 5.6e-200;  
 Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 11 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTGGGCCACGGCGGCGCCCTGG 70  
 Db 450 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTGGGCCACGGCGGCGCCCTGG 509  
 Oy 71 GACCAAGGTGAGACACCCCGTACCTAAABATGAAGGCTGGGGTTGGCTGGCCCTG 130  
 Db 510 GAGCAGAGGTGAGACACCCCGTACCTAAABATGAAGGCTGGGGTTGGCTGGCCCTG 569  
 Oy 131 CTCTGGGGGCGCCCTGCTGGGAAACCCGCTGCGAGAGGAGCAGATCTCCACTGTGA 190  
 Db 570 CTCTGGGGGCGCCCTGCTGGGAAACCCGCTGCGAGAGGAGCAGATCTCCACTGTGA 629  
 Oy 191 GCATCAGAGGCTCTGTGGATGAATAGAAATGGGAAATTTGCCAGGTGAGACCCCAAG 250  
 Db 630 GCATCAGAGGCTCTGTGGATGAATAGAAATGGGAAATTTGCCAGGTGAGACCCCAAG 689  
 Oy 251 ACCATTGATGAGGATCTTTCGCGATCAATCCAGATGAGAGGAGCAGATCTGAGAGTG 310  
 Db 690 ACCATTGATGAGGATCTTTCGCGATCAATCCAGATGAGAGGAGCAGATCTGAGAGTG 749  
 Oy 311 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 370.  
 Db 750 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 809  
 Oy 371 AAGGATATGGGGAGACATTTGATCTTCCACCCATCCAGAAATAGTACTGCTAGTG 430  
 Db 810 AAGGATATGGGGAGACATTTGATCTTCCACCCATCCAGAAATAGTACTGCTAGTG 869  
 Oy 431 GGCCGGAATGGAGATCCAGTGAATGAGACTTACAAAGGATCCGATCGACTCAGATATT 490  
 Db 870 GGCCGGAATGGAGATCCAGTGAATGAGACTTACAAAGGATCCGATCGACTCAGATATT 929  
 Oy 491 AGCGCACCTTCAGATTGCTGCTGAGAGCATTTGTGAGAGAAATACGAGAGTGAATCTCAT 550  
 Db 930 AGCGCACCTTCAGATTGCTGCTGAGAGCATTTGTGAGAGAAATACGAGAGTGAATCTCAT 989  
 Oy 551 GAATCTTTTCCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGGATGAAGGAGACAT 610  
 Db 990 GAATCTTTTCCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGGATGAAGGAGACAT 1049  
 Oy 611 CTCTGTGACATGCGCTCAGATATCGATGATGATGATGATGATGATGATGATGATGAT 670  
 Db 1090 CTCTGTGACATGCGCTCAGATATCGATGATGATGATGATGATGATGATGATGATGAT 1109  
 Oy 671 ACTGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730  
 Db 1110 ACTGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730  
 Oy 731 GTTTTACTGAATTAATGAATAAATATGAATCAAAAGT 771  
 Db 1170 GTTTTACTGAATTAATGAATAAATATGAATCAAAAGT 1210

RESULT 10

US-10-140-474-173

; Sequence 173, Application US/10140474  
 ; Publication No. US20030032156a1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Mei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gunney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tamas, Daniel  
 ; APPLICANT: Watanabe, Collin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P330R1C162  
 ; CURRENT APPLICATION NUMBER: US/10/140,474  
 ; CURRENT FILING DATE: 2002-05-06  
 ; Prior Application removed - See Palm or File Wrapper  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 173  
 ; LENGTH: 1210  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-140-474-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
 Best Local Similarity 98.9%; Pred. No. 5.6e-200;  
 Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 11 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTGGGCCACGGCGGCGCCCTGG 70  
 Db 450 CTGGGGCCAAAGTGAATCCAGCGCTGCGCAGCGCTGGGCCACGGCGGCGCCCTGG 509  
 Oy 71 GACCAAGGTGAGACACCCCGTACCTAAABATGAAGGCTGGGGTTGGCTGGCCCTG 130  
 Db 510 GAGCAGAGGTGAGACACCCCGTACCTAAABATGAAGGCTGGGGTTGGCTGGCCCTG 569  
 Oy 131 CTCTGGGGGCGCCCTGCTGGGAAACCCGCTGCGAGAGGAGCAGATCTCCACTGTGA 190  
 Db 570 CTCTGGGGGCGCCCTGCTGGGAAACCCGCTGCGAGAGGAGCAGATCTCCACTGTGA 629  
 Oy 191 GCATCAGAGGCTCTGTGGATGAATAGAAATGGGAAATTTGCCAGGTGAGACCCCAAG 250  
 Db 630 GCATCAGAGGCTCTGTGGATGAATAGAAATGGGAAATTTGCCAGGTGAGACCCCAAG 689  
 Oy 251 ACCATTGATGAGGATCTTTCGCGATCAATCCAGATGAGAGGAGCAGATCTGAGAGTG 310  
 Db 690 ACCATTGATGAGGATCTTTCGCGATCAATCCAGATGAGAGGAGCAGATCTGAGAGTG 749  
 Oy 311 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 370  
 Db 750 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGTGAGAGAGATATGTGACCGGATG 809  
 Oy 431 GGCCGGAATGGAGATCCAGTGAATGAGACTTACAAAGGATCCGATCGACTCAGATATT 490  
 Db 870 GGCCGGAATGGAGATCCAGTGAATGAGACTTACAAAGGATCCGATCGACTCAGATATT 929  
 Oy 491 AGCGCACCTTCAGATTGCTGCTGAGAGCATTTGTGAGAGAAATACGAGAGTGAATCTCAT 550  
 Db 930 AGCGCACCTTCAGATTGCTGCTGAGAGCATTTGTGAGAGAAATACGAGAGTGAATCTCAT 989



```
QY 551 GAATCTTTTCCCGAGAGCTGACAACTTAAAGCAAACTTTCAGTACGAGACAGAT 610
| | | | |
DB 990 GAATCTTTTCCCGAGAGCTGACAACTTAAAGCAAACTTTCAGTACGAGACAGAT 1049
| | | | |
QY 611 CTTTGTGACCAAGCCCTGACATATCGCATATGACATATACCACTGGAGCAAGCCAC 670
| | | | |
DB 1050 CTTTGTGACCAAGCCCTGACATATCGCATATGACATATACCACTGGAGCAAGCCAC 1109
| | | | |
QY 671 ACTGCGTGTGATGATCAACCCCGAGAGGGGAAATGATGCAATGCCCTTTATATATAT 730
| | | | |
DB 1110 ACTGCGTGTGATGATCAACCCCGAGAGGGGAAATGATGCAATGCCCTTTATATATAT 1169
| | | | |
QY 731 GTTTTACTGAATTAATTAAGTGAACCAAAAGT 771
| | | | |
DB 1170 GTTTTACTGAATTAATTAAGTGAACCAAAAGT 1210
| | | | |
```

```
RESULT 11
; Sequence 173, Application US/10142431
; Publication No. US2003036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Fliviaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumanabe, Colin K
; APPLICANT: Watanabe, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173
```

```
Query Match 93.0%; Score 749.4; DB 9; Length 1210;
Best Local Similarity 98.9%; Pred. No. 5, 6e-200;
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 11 CTTGGGCGCAAGTGAAGTCAAGCGCTGCGAGCGCTTGGGCCAGCGCGCGCGCCCTGG 70
| | | | |
DB 450 CTTGGGCGCAAGTGAAGTCAAGCGCTTGGGCCAGCGCGCGCGCGCCCTGG 509
| | | | |
QY 71 GACCAAAAGGTGAGCAACCCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGAGCCCTG 130
| | | | |
DB 510 GACGAGAGGTGAGCGAGCCCAATTAAGGTAAGATGAAGGCTGGGGTGGCTGGAGCCCTG 569
| | | | |
QY 131 CTTTGGGGGGCCCTCTGGGAAACCCGCTGGGCTGGAGAGGCGAGATCTCACTGTGGA 190
| | | | |
DB 570 CTTTGGGGGGCCCTCTGGGAAACCCGCTGGGCTGGAGAGGCGAGATCTCACTGTGGA 629
| | | | |
QY 191 GCATCAGAGGCTCTGTGGATGAATAGATGGAATTTGCCAGGTGAGACCCCAAGAG 250
| | | | |
DB 630 GCATCAGAGGCTCTGTGGATGAATAGATGGAATTTGCCAGGTGAGACCCCAAGAG 689
| | | | |
```

```
QY 251 ACATTCAGATGGGATCTTCCGATATCATTCAGATGCGACCCAGTCACTGTGAGAGTG 310
| | | | |
DB 690 ACATTCAGATGGGATCTTCCGATATCATTCAGATGCGACCCAGTCACTGTGAGAGTG 749
| | | | |
QY 311 CTTTATGCCCGCTCAGAGGCCACCTCAGAGAGCTGCTGGAGGAGATATGACCCGAGT 370
| | | | |
DB 750 CTTTATGCCCGCTCAGAGGCCACCTCAGAGAGCTGCTGGAGGAGATATGACCCGAGT 809
| | | | |
QY 371 AAGAGATATGGGAAACAGATGATCTTCCAGCCATCCGAAGAACTACAGTACGTAGTG 430
| | | | |
DB 810 AAGAGATATGGGAAACAGATGATCTTCCAGCCATCCGAAGAACTACAGTACGTAGTG 869
| | | | |
QY 431 GCGCGAATGAGAAATCCAGTGAATGAGACTCAAGGATCCGATCCGATCCAGTATAT 490
| | | | |
DB 870 GCGCGAATGAGAAATCCAGTGAATGAGACTCAAGGATCCGATCCGATCCAGTATAT 929
| | | | |
QY 491 AGCGGACCCCTCAAGTTGGCTGTGAGAGCATTTGTGAGAGATACGAGTGAACATAT 550
| | | | |
DB 930 AGCGGACCCCTCAAGTTGGCTGTGAGAGCATTTGTGAGAGATACGAGTGAACATAT 989
| | | | |
QY 551 GAATCTTTTCCCGAGAGCTGACATGTTTAAAGCAAACTTTGACATAGCGAGACAGAT 610
| | | | |
DB 990 GAATCTTTTCCCGAGAGCTGACATGTTTAAAGCAAACTTTGACATAGCGAGACAGAT 1049
| | | | |
QY 611 CTTTGTGACCATGCCCTGACATATGCAATGATGACATGATGACATGACATGACATG 670
| | | | |
DB 1050 CTTTGTGACCATGCCCTGACATATGCAATGATGACATGATGACATGACATGACATG 1109
| | | | |
QY 671 ACTGCGTGTGATGATCAACCCCGAGAGGGGAAATGATGCAATGCCCTTTATATATAT 730
| | | | |
DB 1110 ACTGCGTGTGATGATCAACCCCGAGAGGGGAAATGATGCAATGCCCTTTATATATAT 1169
| | | | |
QY 731 GTTTTACTGAATTAATTAAGTGAACCAAAAGT 771
| | | | |
DB 1170 GTTTTACTGAATTAATTAAGTGAACCAAAAGT 1210
| | | | |
```

```
RESULT 12
; Sequence 173, Application US/10143114
; Publication No. US2003036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Fliviaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumanabe, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-114-173
```

```
Query Match 93.0%; Score 749.4; DB 9; Length 1210;
```

Best Local Similarity 98.9%; Pred. No. 5.6e-200;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 11 CTGGGGCCAAAGTGAAGATCCAGCGCTGCGCCGCTTGGCCAGCGGGGCGCCCTGG 70  
DB 450 CTGGGGCCAAAGTGAAGATCCAGCGCTGCGCCGCTTGGCCAGCGGGGCGCCCTGG 509

OY 71 GACCAAGGTGGAGCAACCCCGTACCTTAARATGAAAGCGTGGGCTTGCTGGCCCTG 130  
DB 510 GAGCAGAGGTGGAGCGACCCCATATAGCTTAAGATGAAAGGCTGGGGTGTGGTGGCCCTG 569

OY 131 CTTCTGGGGGCGCTGCTGGGAACCCGCTGGGCTGCGAGAGCGCAGATCTCCACTGTGA 190  
DB 570 CTTCTGGGGGCGCTGCTGGGAACCCGCTGGGCTGCGAGAGCGCAGATCTCCACTGTGA 629

OY 191 GCATCAGAGGCTCTGTGGATGAACTAGAAATGGGAAATTTGCCAGGTGGACCCCAAGAG 250  
DB 630 GCATCAGAGGCTCTGTGGATGAACTAGAAATGGGAAATTTGCCAGGTGGACCCCAAGAG 689

OY 251 ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAAGTGGAGGTG 310  
DB 690 ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAAGTGGAGGTG 749

OY 311 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGCTGGAGAGATATGTACCCGGATG 370  
DB 750 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGCTGGAGAGATATGTACCCGGATG 809

OY 371 AAGGATATGGGGAAACAGATTGATCTTCCACCCATCCGAAGAACTAGTACGTAGTG 430  
DB 810 AAGGATATGGGGAAACAGATTGATCTTCCACCCATCCGAAGAACTAGTACGTAGTG 869

OY 431 GGCCGGAATGGGAATCCAGTGAACCTGACCTCAAGGCATCCGAATGACTCAGATAT 490  
DB 870 GGCCGGAATGGGAATCCAGTGAACCTGACCTCAAGGCATCCGAATGACTCAGATAT 929

OY 491 AGCGCACCCCTCAAGTTTGCCTGTGAGAGCATTTGAGGAATAGAGGATGAATCAAT 550  
DB 930 AGCGCACCCCTCAAGTTTGCCTGTGAGAGCATTTGAGGAATAGAGGATGAATCAAT 989

OY 551 GAATCTTTTCCGAGAGGCTGACATGTAAAGACAACTTTGACATGAAGGAAACAGAT 610  
DB 990 GAATCTTTTCCGAGAGGCTGACATGTAAAGACAACTTTGACATGAAGGAAACAGAT 1049

OY 611 CTTTGTGACATGCGCTGACATATCGCATGATGAGCTATGAAACACCTGGAGACCCAC 670  
DB 1050 CTTTGTGACATGCGCTGACATATCGCATGATGAGCTATGAAACACCTGGAGACCCAC 1109

OY 671 ACTGCTTGATGATGATCACCCTCCAGAGGAGGAAATGTGGCAATGCTTTATATATAT 730  
DB 1110 ACTGCTTGATGATGATCACCCTCCAGAGGAGGAAATGTGGCAATGCTTTATATATAT 1169

OY 731 GTTTTACTGAATTAATGAAAAAATATGAACCAAAAGT 771  
DB 1170 GTTTTACTGAATTAATGAAAAAATATGAACCAAAAGT 1210

RESULT 13  
US-10-140-002-173  
; Sequence 173, Application US/10140002  
; Publication No. US20030037623A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gettisen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tunas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zhenli  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C59  
; CURRENT APPLICATION NUMBER: US/10/140\_002  
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-002-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 5.6e-200;  
Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 11 CTGGGGCCAAAGTGAAGATCCAGCGCTGCGCCGCTTGGCCAGCGGGGCGCCCTGG 70  
DB 450 CTGGGGCCAAAGTGAAGATCCAGCGCTGCGCCGCTTGGCCAGCGGGGCGCCCTGG 509

OY 71 GACCAAGGTGGAGCAACCCCGTACCTTAARATGAAAGCGTGGGCTTGCTGGCCCTG 130  
DB 510 GAGCAGAGGTGGAGCGACCCCATATAGCTTAAGATGAAAGGCTGGGGTGTGGTGGCCCTG 569

OY 131 CTTCTGGGGGCGCTGCTGGGAACCCGCTGGGCTGCGAGAGCGCAGATCTCCACTGTGA 190  
DB 570 CTTCTGGGGGCGCTGCTGGGAACCCGCTGGGCTGCGAGAGCGCAGATCTCCACTGTGA 629

OY 191 GCATCAGAGGCTCTGTGGATGAACTAGAAATGGGAAATTTGCCAGGTGGACCCCAAGAG 250  
DB 630 GCATCAGAGGCTCTGTGGATGAACTAGAAATGGGAAATTTGCCAGGTGGACCCCAAGAG 689

OY 251 ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAAGTGGAGGTG 310  
DB 690 ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAAGTGGAGGTG 749

OY 311 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGCTGGAGAGATATGTACCCGGATG 370  
DB 750 CCTTATGCGCGCTCAGAGGCCCACTCAAGAGCTGCTGGAGAGATATGTACCCGGATG 809

OY 371 AAGGATATGGGGAAACAGATTGATCTTCCACCCATCCGAAGAACTAGTACGTAGTG 430  
DB 810 AAGGATATGGGGAAACAGATTGATCTTCCACCCATCCGAAGAACTAGTACGTAGTG 869

OY 431 GGCCGGAATGGGAATCCAGTGAACCTGACCTCAAGGCATCCGAATGACTCAGATAT 490  
DB 870 GGCCGGAATGGGAATCCAGTGAACCTGACCTCAAGGCATCCGAATGACTCAGATAT 929

OY 491 AGCGCACCCCTCAAGTTTGCCTGTGAGAGCATTTGAGGAATAGAGGATGAATCAAT 550  
DB 930 AGCGCACCCCTCAAGTTTGCCTGTGAGAGCATTTGAGGAATAGAGGATGAATCAAT 989

OY 551 GAATCTTTTCCGAGAGGCTGACATGTAAAGACAACTTTGACATGAAGGAAACAGAT 610  
DB 990 GAATCTTTTCCGAGAGGCTGACATGTAAAGACAACTTTGACATGAAGGAAACAGAT 1049

OY 611 CTTTGTGACATGCGCTGACATATCGCATGATGAGCTATGAAACACCTGGAGACCCAC 670  
DB 1050 CTTTGTGACATGCGCTGACATATCGCATGATGAGCTATGAAACACCTGGAGACCCAC 1109

OY 671 ACTGCTTGATGATGATCACCCTCCAGAGGAGGAAATGTGGCAATGCTTTATATATAT 730  
DB 1110 ACTGCTTGATGATGATCACCCTCCAGAGGAGGAAATGTGGCAATGCTTTATATATAT 1169

OY 731 GTTTTACTGAATTAATGAAAAAATATGAACCAAAAGT 771  
DB 1170 GTTTTACTGAATTAATGAAAAAATATGAACCAAAAGT 1210

RESULT 14  
US-10-142-419-173

Sequence 173, Application US/10142419  
Publication No. US20030044945A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C244  
CURRENT APPLICATION NUMBER: US/10/142, 419  
CURRENT FILING DATE: 2002-05-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-142-419-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 5, 6e-200;

Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCAAAGTGAAGTCAAGGCTGCTGCAAGCGTTGGGCGACGGGGCCCTGG 70  
450 CTGGGGCCAAAGTGAAGTCAAGGCTGCTGCAAGCGTTGGGCGACGGGGCCCTGG 509  
71 GACCAAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGCCCTG 130  
510 GACGAGAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGCCCTG 569  
131 CTCTGGGGCCCTGCTGGGGAACCGCTGGGCTGGGAGGAGCCAGATCTCCACTGTGA 190  
570 CTCTGGGGCCCTGCTGGGGAACCGCTGGGCTGGGAGGAGCCAGATCTCCACTGTGA 629  
191 GCATGAGGGGCTGCTGGTGAAGTGAATGGAATTCGCCAGGAGGAGCCCAAGAG 250  
630 GCATGAGGGGCTGCTGGTGAAGTGAATGGAATTCGCCAGGAGGAGCCCAAGAG 689  
251 ACATTCAGATGGGATCTTCCGATTCATCCAGATGGAGCCAGCTCAGTGGAGGTG 310  
690 ACATTCAGATGGGATCTTCCGATTCATCCAGATGGAGCCAGCTCAGTGGAGGTG 749  
311 CCTATGCGCGCTCAGAGGCGCACTCAGAGCTGCTGGAGAGATATGTGACCGGATG 370  
750 CCTATGCGCGCTCAGAGGCGCACTCAGAGCTGCTGGAGAGATATGTGACCGGATG 809  
371 AAGAGATATGGGAGACAGATGATCTTCCAGCCATCGCAAGAACTACGTAAGTGTAGT 430  
810 AAGAGATATGGGAGACAGATGATCTTCCAGCCATCGCAAGAACTACGTAAGTGTAGT 869  
431 GGGCGAATGGAGAAATTCAGTGAATCGACTCTCAAGGCAATCCGAATCGACTGATAT 490  
870 GGGCGAATGGAGAAATTCAGTGAATCGACTCTCAAGGCAATCCGAATCGACTGATAT 929

491 AGCGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGTGGAGAAATAGAGGATGACCTATT 550  
930 AGCGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGTGGAGAAATAGAGGATGACCTATT 989  
551 GAATCTTTTCCGAGAGGCTGACATGTTAAAGACAACTTTGACAGTAAGCAAGAT 610  
990 GAATCTTTTCCGAGAGGCTGACATGTTAAAGACAACTTTGACAGTAAGCAAGAT 1049  
611 CTCTGGAGCATGCCCGCAATATCCGATGATGAGTATGAGACCATGAGAGCCAC 670  
1050 CTCTGGAGCATGCCCGCAATATCCGATGATGAGTATGAGACCATGAGAGCCAC 1109  
671 ACTGCTTGAATGATACACCCGAGAGGAGGAAATGGTGCATGCTTTATATATAT 730  
1110 ACTGCTTGAATGATACACCCGAGAGGAGGAAATGGTGCATGCTTTATATATAT 1169  
731 GTTTTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 771  
1170 GTTTTCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1210

RESULT 15  
US-10-123-262-173

Sequence 173, Application US/10123262  
Publication No. US20030049816A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C38  
CURRENT APPLICATION NUMBER: US/10/123, 262  
CURRENT FILING DATE: 2002-04-15  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-123-262-173

Query Match 93.0%; Score 749.4; DB 9; Length 1210;  
Best Local Similarity 98.9%; Pred. No. 5, 6e-200;

Matches 753; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

11 CTGGGGCCAAAGTGAAGTCAAGGCTGCTGCAAGCGTTGGGCGACGGGGCCCTGG 70  
450 CTGGGGCCAAAGTGAAGTCAAGGCTGCTGCAAGCGTTGGGCGACGGGGCCCTGG 509  
71 GACCAAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGCCCTG 130  
510 GACGAGAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGGTGGCTGGCCCTG 569  
131 CTCTGGGGCCCTGCTGGGGAACCGCTGGGCTGGGAGGAGCCAGAGTCTCCACTGTGA 190  
570 CTCTGGGGCCCTGCTGGGGAACCGCTGGGCTGGGAGGAGCCAGAGTCTCCACTGTGA 629  
191 GCATGAGGGGCTGCTGGTGAATGAATGGAATTCGCCAGGTGAGCCCAAGAG 250

```

Db      630 GCGATGAGGGCTCTGGTGGATGAACAGATGGAAATGGCCAGGTGGACCCCAAGAAG 689
QY      251 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTCAGTGGTAGTG 310
Db      690 ACCATTGAGATGGATCTTTCCGATCAATCCAGATGGCAGCCAGTCAGTGGTAGTG 749
QY      311 CCTTATGCCCGCTCAGAGGCCCACTCAGACAGCTGCTGGAGAGATATGTACCGGATG 370
Db      750 CCTTATGCCCGCTCAGAGGCCCACTCAGACAGCTGCTGGAGAGATATGTACCGGATG 809
QY      371 AAGGAGTATGGGAACAGATGATCCTTCCACCATCCGACAGAACATGACGTAGTGTG 430
Db      810 AAGGAGTATGGGAACAGATGATCCTTCCACCATCCGACAGAACATGACGTAGTGTG 869
QY      431 GCGCGAATGAGAAATCCAGTAACTGAGACTACAAGGCATCCGAATCGACTCAGATATT 490
Db      870 GCGCGAATGAGAAATCCAGTAACTGAGACTACAAGGCATCCGAATCGACTCAGATATT 929
QY      491 AGCGGCACCCCTCAAGTTGCGTGTGAGAGCAATGTGTGAGGAATACGAGGATGAACATCATT 550
Db      930 AGCGGCACCCCTCAAGTTGCGTGTGAGAGCAATGTGTGAGGAATACGAGGATGAACATCATT 989
QY      551 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTGCAGTAAAGCAACAGAT 610
Db      990 GAATCTTTTCCGAGAGGCTGACAAATGTTAAAGACAACTTGCAGTAAAGCAACAGAT 1049
QY      611 CTTTGTGACCATGCGCTGCACATATCGCATGTGAGCTATGAACCACTGAGCAGGCCAC 670
Db      1050 CTTTGTGACCATGCGCTGCACATATCGCATGTGAGCTATGAACCACTGAGCAGGCCAC 1109
QY      671 ACTGGCTTATGATGATCACCCTCCAGAGGGGAAATGTGTGCAATGCTTTTATATATTAT 730
Db      1110 ACTGGCTTATGATGATCACCCTCCAGAGGGGAAATGTGTGCAATGCTTTTATATATTAT 1169
QY      731 GTTTTACTGAATTTACTGAAAAAATATGAAACCAAAAGT 771
Db      1170 GTTTTACTGAATTTACTGAAAAAATATGAAACCAAAAGT 1210

```

Search completed: March 30, 2003, 22:23:35  
 Job time : 86.3669 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:36:03 ; Search time 1147.74 Seconds

(without alignments)  
11373.235 Million cell updates/sec

Title: US-10-082-502-16

Perfect score: 806

Sequence: 1 cggcccaagcctggggccca.....aagagagagagagaacta 806

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	731	90.7	1117	13	BM552907
2	714.2	88.6	783	13	BM550436
3	713.2	88.5	868	14	BM550436
4	705.4	87.5	880	14	BM550436
5	698.6	86.7	875	13	BM550436
6	693	86.0	773	13	BM550162

7	676.6	83.9	742	12	BM524549	BM524549	602728454
8	674.4	83.3	697	12	BM544334	BM544334	602014708
9	674.4	83.3	697	12	BM544334	BM544334	602014708
10	669.8	83.1	673	14	BM668530	BM668530	603048045
11	660.8	82.0	720	13	BM190975	BM190975	603070841
12	650.8	80.7	914	12	BM799495	BM799495	601589289
13	646.4	80.2	791	12	BM704443	BM704443	602688633
14	645	80.0	765	12	BM712281	BM712281	601898804
15	641	79.5	745	12	BM791763	BM791763	601898804
16	639	79.3	818	13	BM870315	BM870315	601447464
17	636.8	79.0	783	12	BM763784	BM763784	603049729
18	628.2	77.9	719	12	BM733900	BM733900	601568370
19	623.8	77.4	746	13	BM551339	BM551339	603193144
20	620	76.9	856	12	BM965088	BM965088	602268856
21	616.6	76.5	751	12	BM742698	BM742698	601574753
22	614.6	76.3	699	9	AL548119	AL548119	601574753
23	607.8	75.4	530	14	BM437660	BM437660	602728454
24	604.2	75.0	611	13	BM171899	BM171899	602728454
25	603.6	74.9	853	12	BM403049	BM403049	602728454
26	602.2	74.7	797	13	BM762927	BM762927	603047987
27	599	74.3	599	13	BM791489	BM791489	602728454
28	590.6	73.3	591	14	BM632888	BM632888	602728454
29	589.8	73.2	593	10	AM167673	AM167673	602728454
30	587.6	72.9	588	14	BM832828	BM832828	602728454
31	586.8	72.8	885	14	BM920767	BM920767	602728454
32	586.8	72.8	963	14	BM920767	BM920767	602728454
33	585.4	72.6	1019	12	BM920767	BM920767	602728454
34	584.4	72.5	908	14	BM087139	BM087139	602728454
35	583.8	72.4	1255	11	AK019927	AK019927	602728454
36	583.2	72.4	772	11	AK007914	AK007914	602728454
37	583	72.3	892	13	BM523169	BM523169	601175203
38	580.4	72.0	776	13	BM523169	BM523169	601175203
39	578	71.7	587	10	BE042694	BE042694	602728454
40	577.2	71.6	754	11	AK013568	AK013568	602728454
41	574	71.2	853	11	AK013014	AK013014	602728454
42	572	71.0	594	14	BM020508	BM020508	602728454
43	571.6	70.9	587	14	AI860382	AI860382	602728454
44	570.4	70.8	808	13	BI408525	BI408525	602728454
45	570	70.7	598	9	AI418940	AI418940	602728454

#### ALIGNMENTS

RESULT 1  
LOCUS BM552907 1117 bp mRNA linear EST 20-FEB-2002  
DEFINITION AGENCOURT.6572552 NIH\_MGC\_41 Homo sapiens CDNA clone IMAGE:5466942  
5' mRNA sequence.  
ACCESSION BM552907  
VERSION BM552907.1 GI:18791172  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1117)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LCM1968 row: C column: 07  
High quality sequence stop: 709.  
Location/Qualifiers 1. 1117

#### FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5466942"
/clone_1lb="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT      299 a      258 c      320 g      239 t      1 others
ORIGIN

```

```

Query Match      90.7%; Score 731; DB 13; Length 1117;
Best Local Similarity 95.5%; Pred. No. 5e-119;
Matches 762; Conservative 1; Mismatches 32; Indels 3; Gaps 1;

```

```

OY 11 CTGGGCGCAAGTGAAGTCCAGCGCTGCTGCCAGCGCTGGGCGCACGGCGGCGCTGCG 70
DB 55 CTGGGCGCAAGTGAAGTCCAGCGCTGCTGCCAGCGCTGGGCGCACGGCGGCGCTGCG 114
OY 71 GACCAAGGTGAGCAACCCGCTTACCTTAARATGAAGGCTGGGCTGGCTGCCCTG 130
DB 115 GAGCAGAGGTGAGCGACCCCTTACGCTAAAGATGAAGGCTGGGCTGGCTGCCCTG 174
OY 131 CTTGCGGGGGCGCTGCTGGGAACCGCGCTGGGCTGGAGAGCGCAGATCTCCACTGTGA 190
DB 175 CTTGCGGGGGCGCTGCTGGGAACCGCGCTGGGCTGGAGAGCGCAGATCTCCACTGTGA 234
OY 191 GCATGCGAGGCTGCTGGTGAATGAATGAATGGAAATTCGCCAGGTGAGCCCAAGAAG 250
DB 235 GCATGCGAGGCTGCTGGTGAATGAATGAATGGAAATTCGCCAGGTGAGCCCAAGAAG 294
OY 251 ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGCGAGCAGTCACTGGTGAAGTG 310
DB 295 ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGCGAGCAGTCACTGGTGAAGTG 354
OY 311 CCTTATGCGCGCTGAGAGCGCCACTCAAGAGTGTGAGAGAGATATGTGACCGGATG 370
DB 355 CCTTATGCGCGCTGAGAGCGCCACTCAAGAGTGTGAGAGAGATATGTGACCGGATG 414
OY 371 AAGGATGAGGGAACAGATTGATCTTCCACCATTCGCAAGAACTACGTAGTGTAGTG 430
DB 415 AAGGATGAGGGAACAGATTGATCTTCCACCATTCGCAAGAACTACGTAGTGTAGTG 474
OY 431 GCGCGGATGGAATCAAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGT 490
DB 475 GCGCGGATGGAATCAAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGT 534
OY 491 AGCGGACCCCTCAAGTGTGCTGAGAGCAATTTGAGAGAAATAGAGAGTGAATGATCT 550
DB 535 AGCGGACCCCTCAAGTGTGCTGAGAGCAATTTGAGAGAAATAGAGAGTGAATGATCT 594
OY 551 GAATTTCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGCAATGAGCAAGAT 610
DB 595 GAATTTCTTTCCGAGAGGCTGACAAATGTTAAAGCAAACTTTGCAATGAGCAAGAT 654
OY 611 CTTTGTGACCATGCGCTGCAATATCGATGATGAGTATGACCATGAGAGAGCCAC 670
DB 655 CTTTGTGACCATGCGCTGCAATATCGATGATGAGTATGACCATGAGAGAGCCAC 714
OY 671 ACTGCGTGAATGATCAACCCAGAGAGGGAATGATGAGCAATGAGC 727
DB 715 ACTGCGTGAATGATCAACCCAGAGAGGGAATGATGAGCAATGAGC 774
OY 728 TATGTTTTTACTGAATTAATGAATAATGAATGAATGAATGAATGAATGAATGAATGAAT 787
DB 775 TGTGTTTTTACTGAATTAATGAATAATGAATGAATGAATGAATGAATGAATGAATGAAT 834

```

```

OY 788 AGAGAGAGAGAGAGACT 805
DB 835 AAAAACTCGAAGCATTT 852

```

```

RESULT 2
LOCUS      BI550436
DEFINITION 603192905P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5264328 5',
            mRNA sequence.
ACCESSION  BI550436
VERSION    BI550436.1 GI:15437748
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 783)
            NIH-MGC http://mhc.ncl.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL  Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgrsbs-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraiki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: LNL11666 row: 1 column: 01
            High quality sequence stop: 730.
            Location/Qualifiers

```

```

FEATURES
    source          1..783
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:5264328"
                    /clone_1lb="NIH_MGC_95"
                    /tissue_type="hippocampus"
                    /lab_host="DH10B"
                    /note="Organ: brain; Vector: pBluescript (modified
                    pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                    ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
                    size-selected for average insert size 2.5 kb and
                    normalized to ROP 5. This is a primary library enriched
                    for full-length clones and constructed using the
                    Cap-trapper method (Carninci, in preparation). Library
                    constructed by M. Brownstein (NIH/NHGRI, National
                    Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      209 a      189 c      225 g      160 t
ORIGIN

```

```

Query Match      88.6%; Score 714.2; DB 13; Length 783;
Best Local Similarity 97.1%; Pred. No. 5.3e-116;
Matches 747; Conservative 1; Mismatches 19; Indels 2; Gaps 2;

```

```

OY 21 AGTGAAGTCCAGCGCTGCGCAGCGCTTGGGCGCACGGCGGCGCTGGACCAAGAT 80
DB 1 AGCGGGAATCCAGCGCTTCCAGCGCTTGGGCGCACGGCGGCGCTGGGAGCAGAGT 60
OY 81 GGAGCAACCCGCTTACCTTAARATGAAGGCTGGGCTGGGCTGCTGCTGCGGGG 140
DB 61 GGAGGACCCCATTTACGCTAAAGATGAAGGCTGGGCTGGGCTGCTGCTGCGGGG 120
OY 141 CCCCTGCGGAACCGCGCTGGGCTGGAGAGCAGAGATCTCCACTGTGAGACATGCAAGG 200
DB 121 CCCCTGCGGAACCGCGCTGGGCTGGAGAGCAGAGATCTCCACTGTGAGACATGCAAGG 180
OY 201 CTCGTGATGATGAATGAATGGAAATTTGCCAGGTGAGACCCCAAGAACCATTCAGA 260
DB 181 CTCGTGATGATGAATGAATGGAAATTTGCCAGGTGAGACCCCAAGAACCATTCAGA 240

```

Oy 261 TGGGATCTTCCGATCAATCCAGATGGACACCACTAGTGTGGAGTGCTTATGCC 320  
 Db 241 TGGGATCTTCCGATCAATCCAGATGGACACCACTAGTGTGGAGTGCTTATGCC 300  
 Oy 321 GCTCAGAGGCCACCTCAGAGAGCTGCTGAGAGATATGAGCCGATGAAGAGATG 380  
 Db 301 GCTAAGAGGCCACCTCAGAGAGCTGCTGAGAGATATGAGCCGATGAAGAGATG 360  
 Oy 381 GGAAGAGATGATCTTCCAGCCATCGCAAGAACTACGATCTGATGAGGCGGATG 440  
 Db 361 GGAAGAGATGATCTTCCAGCCATCGCAAGAACTACGATCTGATGAGGCGGATG 420  
 Oy 441 GAGAATCAGTGAAGTGAAGTGAAGGATCCCAATCGACTGATATTAAGCCACCC 500  
 Db 421 GAGAATCAGTGAAGTGAAGTGAAGGATCCCAATCGACTGATATTAAGCCACCC 480  
 Oy 501 TCAAGTTGGGTGAGAGATGAGAGATACGAGATGAATGATGATGATGATGATG 560  
 Db 481 TCAAGTTGGGTGAGAGATGAGAGATACGAGATGAATGATGATGATGATGATG 540  
 Oy 561 CCCGAGAGGCTGACATGTTAAAGACAACCTTGGCAAGCGAAGATCTTGTGACC 620  
 Db 541 CCCGAGAGGCTGACATGTTAAAGACAACCTTGGCAAGCGAAGATCTTGTGACC 600  
 Oy 621 ATGCGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 680  
 Db 601 ATGCGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 659  
 Oy 681 TGATCAGCCCCAGAGAGGAGAAATGTTGGCAATGCC-TTTTAAATATATGTTTACT 739  
 Db 660 TGATCAGCCCCAGAGAGGAGAAATGTTGGCAATGCC-TTTTAAATATATGTTTACT 719  
 Oy 740 GAATTTACGAAAAATATGAACCAAAAGTAAAAAATAAAAA 788  
 Db 720 GAATTTACGAAAAATATGAACCAAAAGTAAAAAATAAAAA 768

RESULT 3  
 LOCUS BQ421069 868 bp mRNA linear EST 23-MAY-2002  
 DEFINITION AGENCOURT\_7911956 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6009949  
 5', mRNA sequence.  
 ACCESSION BQ421069  
 VERSION BQ421069.1 GI:21116384  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NIH-MGC http://mhc.nci.nih.gov/  
 1 (bases 1 to 868)  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: DCTD/DP/GenZar  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNI)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LILNI at:  
 http://image.llnl.gov  
 Plate: LILNI13197 row: d column: 14  
 High quality sequence stop: 542.  
 Location/Qualifiers  
 1..868  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6009949"  
 /clone\_1="NIH\_MGC\_68"  
 /tissue-type="large cell carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: PCMV-SPORE6; Site\_1: NCI;"

BASE COUNT 232 a 203 c 257 g 168 t 8 others  
 ORIGIN  
 Query Match 88.5%; Score 713.2; DB 14; Length 868;  
 Best Local Similarity 94.6%; Pred. No. 7 6e-116;  
 Matches 755; Conservative 1; Mismatches 37; Indels 5; Gaps 2;  
 Oy 11 CTGGGCCCCAAGTGAAGTCCAGCGGTCTGCGAGCCGCTGGGCGCAGCGGCCCTGG 70  
 Db 37 CTGGGCCCCAAGTGAAGTCCAGCGGTCTGCGAGCCGCTGGGCGCAGCGGCCCTGG 96  
 Oy 71 GACCAAGGTGGAGCAACCCGTTACCTTAAATGAAAGCTGGGTTGGTGGCCCTG 130  
 Db 97 GAGCAGAGGTGGAGCAACCCGTTACCTTAAATGAAAGCTGGGTTGGTGGCCCTG 156  
 Oy 131 CTCTGGGGCCCTGCTGGGAAACCGCTGGGCTCGAGAGCCAGATCTCACTGGGA 190  
 Db 157 CTCTGGGGCCCTGCTGGGAAACCGCTGGGCTCGAGAGCCAGATCTCACTGGGA 216  
 Oy 191 GCATGCAAGGCTCTGGTGAATGAATGGAATGGCAATGGCCAGGTGACCCCAAGAG 250  
 Db 217 GCATGCAAGGCTCTGGTGAATGAATGGAATGGCAATGGCCAGGTGACCCCAAGAG 276  
 Oy 251 ACCATTGAGTGGATCTTCCGGATCAATCCAGATGGCCAGGATGATGATGATGATG 310  
 Db 277 ACCATTGAGTGGATCTTCCGGATCAATCCAGATGGCCAGGATGATGATGATGATG 336  
 Oy 311 CCTTATGCCGCTCAGAGGCCACCTCACAGACCTCTGGAGAGATATGATGATGATG 370  
 Db 337 CCTTATGCCGCTCAGAGGCCACCTCACAGACCTCTGGAGAGATATGATGATGATG 396  
 Oy 371 AAGAGATATGGGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 430  
 Db 397 AAGAGATATGGGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 456  
 Oy 431 GGCGGAATGAGAGATCCAGTGAACCTGACCTCAAGGATCCGATCCAGTCAATGAT 490  
 Db 457 GGCGGAATGAGAGATCCAGTGAACCTGACCTCAAGGATCCGATCCAGTCAATGAT 516  
 Oy 491 AGCGGCAACCTCAAGTTCGCTGAGAGCATTTGGAGAAATACGAGATGAATCAT 550  
 Db 517 AGCGGCAACCTCAAGTTCGCTGAGAGCATTTGGAGAAATACGAGATGAATCAT 576  
 Oy 551 GAATTTCTTCCGAGAGGCTGACATGTTAAAGACAACTTTGACGTAAAGCAAGAT 610  
 Db 577 GAATTTCTTCCGAGAGGCTGACATGTTAAAGACAACTTTGACGTAAAGCAAGAT 636  
 Oy 611 CTTTGTGACCATGCTCTGACATATGATGATGATGATGATGATGATGATGATGATG 670  
 Db 637 CTTTGTGACCATGCTCTGACATATGATGATGATGATGATGATGATGATGATGATG 696  
 Oy 671 ACTGCTTATGATGATCACCCTCAGAGAGGGAAT--GGTGGCAATGCTTTATATAT 728  
 Db 697 ACTGCTTATGATGATCACCCTCAGAGAGGGAAT--GGTGGCAATGCTTTATATAT 756  
 Oy 729 ATG--TTTCTGTAATTAATGAAAAATATGAACCAAAAGTAAAAAATAAAAA 785  
 Db 757 ATGTTTCTTACCGGAATTAATGAAAAATATGAACCAAAAGTAAAAAATAAAAA 816  
 Oy 786 AAAGAGAGAGAGAGAGAA 803  
 Db 817 AAAGAGAGAGAGAGAGAA 834

RESULT 4  
 LOCUS BQ421827 880 bp mRNA linear EST 23-MAY-2002  
 DEFINITION AGENCOURT\_7801623 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6050568  
 5', mRNA sequence.  
 ACCESSION BQ421827



```

VERSION      B0421827.1  GI:21117142
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 880)
              NIH-MGC http://mgc.ncl.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              JOURNAL
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs-femail.nih.gov
              Tissue Procurement: ATCC/DCTD/DFP
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://lmgc.lnl.gov
              Plate: LLM11303 row: a column: 01
              High quality sequence stop: 580.
FEATURES     location/Qualifiers
              1..880
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:6050568"
              /clone_11b="NIH-MGC_72"
              /tissue_type="melanocytic melanoma"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: skin; Vector: pCMV-SPOB6; Site_1: NctI;
              Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.
              Average insert size 2 kb. Library constructed by Life
              Technologies."
BASE COUNT   238 a 208 c 256 g 176 t 2 others
ORIGIN
Query Match      87.5%; Score 705.4; DB 14; Length 880;
Best Local Similarity 96.8%; Pred. No. 1.8e-114;
Matches 760; Conservative 1; Mismatches 19; Indels 5; Gaps 4;
OY 11 CTGGGCGCAAAATGAAATGACAGCGCTGTGCGCGCTTGCGCGCGCGCGCGCTTG 70
Db 9 CTGGGCGCAAAATGAAATGACAGCGCTGTGCGCGCTTGCGCGCGCGCGCGCTTG 68
OY 71 GACCAAGGTGAGCAACCCCGTACCTAAARATGAAGGCTGGGCTTGCGCGCGCTTG 130
Db 69 GAGCAGAGCTGAGCGACCCCACTACGCTAAAGATGAAGGCTGGGCTTGCGCGCGCTTG 128
OY 131 CTTCGGGGGGCCCTGCTGGGGAACCGCGCTGGGCTGGGAGGAGGAGGAGTCTCAGTGTGA 190
Db 129 CTTCGGGGGGCCCTGCTGGGGAACCGCGCTGGGCTGGGAGGAGGAGGAGTCTCAGTGTGA 188
OY 191 GCATGACAGGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 250
Db 189 GCATGACAGGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 248
OY 251 ACCATTGAGATGGGATCTTTCCGGATCAATCCAGATGGCAGGCACTGAGTGTGA 310
Db 249 ACCATTGAGATGGGATCTTTCCGGATCAATCCAGATGGCAGGCACTGAGTGTGA 308
OY 311 CCTTATGCGCGCTGAGAGGCGCACTCAGAGAGTGTGGAGAGATATGTACCGGATG 370
Db 309 CCTTATGCGCGCTGAGAGGCGCACTCAGAGAGTGTGGAGAGATATGTACCGGATG 368
OY 371 AAGGATATGGGGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 430
Db 369 AAGGATATGGGGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 428
OY 431 GGGCGGAATGGAGATCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 490
Db 429 GGGCGGAATGGAGATCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 488
OY 491 AGCGGACACCTCAAGTTTGGCTGAGAGCAATGTGGAGAGATGAGAGATCAACTATT 550

```

```

Db 489 ACCGGACCCCTCAAGTTTGGCTGAGAGCAATGTGGAGAGATGAGAGATCAACTATT 548
OY 551 GAATCTTTCCCGAGAGAGGCTGACATGTTAAAGACAACTTTGCGAGTAAAGACAGAT 610
Db 549 GAATCTTTCCCGAGAGAGGCTGACATGTTAAAGACAACTTTGCGAGTAAAGACAGAT 608
OY 611 CTTTGTGACCATGCGCTGACATATGCGATGATGAGCTATGAAACAGTGGACAGCCAC 670
Db 609 CTTTGTGACCATGCGCTGACATATGCGATGATGAGCTATGAAACAGTGGACAGCCAC 668
OY 671 ACT-GGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
Db 669 ACTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
OY 729 -ATGTTTACTGAAATTAATG--AAAAATATGAAACAAAGTAAAGTAAAGTAAAGTAAAGT 785
Db 729 AATGTTTACTGAAATTAATG--AAAAATATGAAACAAAGTAAAGTAAAGTAAAGTAAAGT 788
OY 786 AAAGA 790
Db 789 AAAAA 793

```

```

RESULT 5
BI763950 875 bp mRNA linear EST 25-SEP-2001
LOCUS 603049829P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5189862 5',
DEFINITION
        mRNA sequence.
ACCESSION
        BI763950
VERSION
        BI763950.1 GI:15755528
KEYWORDS
        EST.
SOURCE
        human.
ORGANISM
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
        1 (bases 1 to 875)
        NIH-MGC http://mgc.ncl.nih.gov/.
        National Institutes of Health, Mammalian Gene Collection (MGC)
        JOURNAL
        Unpublished (1999)
        Contact: Robert Strausberg, Ph.D.
        Email: cga@bbs-femail.nih.gov
        Tissue Procurement: Life Technologies, Inc.
        cDNA Library Preparation: Life Technologies, Inc.
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
        DNA Sequencing by: Incyte Genomics, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LNLN at:
        http://lmgc.lnl.gov
        Plate: LLM11474 row: j column: 07
        High quality sequence start: 3
        High quality sequence stop: 834.
FEATURES     location/Qualifiers
              1..875
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:5189862"
              /clone_11b="NIH-MGC_116"
              /lab_host="DH10B"
              /note="Organ: pooled colon, kidney, stomach; Vector:
              pCMV-SPOB6; Site_1: NctI; Site_2: EcoRV (destroyed); RNA
              source anonymous pool of 3 colons, age 26 yo male, 49 yo
              female, 71 yo male colon; 46 yo male kidney, and pool of 2
              stomachs, 62 yo male and 70 yo female. Library is
              oligo-dT primed and directionally cloned (EcoRV site is
              destroyed upon cloning). Average insert size 1.4 kb,
              insert size range 1-3 kb. Library is normalized and
              enriched for full-length clones and was constructed by C.
              Gruber (Invitrogen). Research Genetics tracking code
              023. Note: this is a NIH-MGC Library."
BASE COUNT   226 a 209 c 259 g 181 t
ORIGIN

```



Db 481 TCAAGTTGGCTGAGACATTGGAGCAATACGAGATCAACATGTAATCTTTT 540  
 QY 561 CCCGAGAGCTGACATGTTAAAGACAACTTGCAGTAAGCAACAGATC-TTTGTGAC 619  
 Db 541 CCCGAGAGCTGACATGTTAAAGCAAACTTGCAGTAAGCAACAGATCTTTGTGAC 600  
 QY 620 CATGCCCTGACATATGCGATGATGAGTATGACCACTGAGCAGCCACACTGGCTTG 679  
 Db 601 CATGCCCTGACATATGCGATGATGAGTATGACCACTGAGCAGCCACACTGGCTTG 660  
 QY 680 ATGGATCAACCCCGAGAGGG--AAAATGGTGCAATGCTTTATATTTATGTTTA 737  
 Db 661 ATGGATCAACCCCGAGAGGGAAATGTTGGCAATGCTTTATATTTATGTTTA 720  
 QY 738 CTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 785  
 Db 721 CTGAATTA-CTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 766

RESULT 7  
 BG824549 742 bp mRNA linear EST 22-MAY-2001  
 LOCUS 602728454F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4868100 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG824549  
 VERSION BG824549.1 GI:14172136  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 742)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: NIH Intramural Sequencing Center  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1CM1737 row: k column: 13  
 High quality sequence stop: 740.

FEATURES  
 Source Location/Qualifiers  
 1..742  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="4868100"  
 /clone\_id="NIH.MGC.15"  
 /tissue\_type="adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: colon; Vector: pOTB1; site\_1: XhoI; site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(C). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)"  
 BASE COUNT 180 a 184 c 233 g 144 t 1 others  
 ORIGIN

Query Match 83.9%; Score 676.6; DB 12; Length 742;  
 Best Local Similarity 98.3%; Pred. No. 2.2e-109;  
 Matches 682; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 11 CTGGGGCCAAAGTGAAGCCAGCGGTGCGAGCGCTTGGCCAGCGGGCGCCCTGG 70  
 Db 48 CTGGGGCCAAAGTGAAGCGGTCTCCAGCGCTTGGCGGCGGGCGCCCTGG 107

QY 71 GACCAAGGTGAGCAACCCGTTACCTTAARATGAAAGGTGGGTTGGTGGCCCTG 130  
 Db 108 GAGCAGAGTGGAGCAACCCCAATTACGTAAGATGAAGAGGTGGGTTGGTGGCCCTG 167  
 QY 131 CTCTCTGGGGCCCTGCTGGGAACCGCTTGGGCTCGGAGAGCAAGATCTCACTGTGA 190  
 Db 168 CTCTCTGGGGCCCTGCTGGGAACCGCTTGGGCTCGGAGAGCAAGATCTCACTGTGA 227  
 QY 191 GCATCAGAGGCTCTGGTGTGATGACATAGAAATGGAAATGGCCAGGTGACCCCAAGAG 250  
 Db 228 GCATCAGAGGCTCTGGTGTGATGACATAGAAATGGAAATGGCCAGGTGACCCCAAGAG 287  
 QY 251 ACCATTACATGGGATCTTCCGGATCAATCCAGATGGCAGCCAGTGTGGAGTG 310  
 Db 288 ACCATTACATGGGATCTTCCGGATCAATCCAGATGGCAGCCAGTGTGGAGTG 347  
 QY 311 CCTTATGCCCCGCTCAGAGGCCCACTCAGAGCTCTGTGGAGAGATATGTAGCGGATG 370  
 Db 348 CCTTATGCCCCGCTCAGAGGCCCACTCAGAGCTCTGTGGAGAGATATGTAGCGGATG 407  
 QY 371 AAGGATATGGGGAACAGATGATCTCTCCACCATCGCAAGACTAGCTAGCTAGTG 430  
 Db 408 AAGGATATGGGGAACAGATGATCTCTCCACCATCGCAAGACTAGCTAGCTAGTG 467  
 QY 431 GGCCGGATGAGAAATCCAGTGAATGACCTGACCTGCAAGGCTCGAATGACTCAGATAT 490  
 Db 468 GGCCGGATGAGAAATCCAGTGAATGACCTGACCTGCAAGGCTCGAATGACTCAGATAT 527  
 QY 491 AGCGGACCCCTCAAGTTTCCGTGTGAGACATTTGGAGAAATAGAGATGAATCTATT 550  
 Db 528 AGCGGACCCCTCAAGTTTCCGTGTGAGACATTTGGAGAAATAGAGATGAATCTATT 587  
 QY 551 GAATCTTTTCCGAGAGGCTGACATATTTAAAGCAAACTTGGAGTAAGCAAGAT 610  
 Db 588 GAATCTTTTCCGAGAGGCTGACATATTTAAAGCAAACTTGGAGTAAGCAAGAT 647  
 QY 611 CTCTGTGACCAATGCCCTGACATATTCGATGATGAGCTATGAACCACTGGACAGCCAC 670  
 Db 648 CTCTGTGACCAATGCCCTGACATATTCGATGATGAGCTATGAACCACTGGACAGCCAC 707  
 QY 671 ACTGCTTATGATGATCAACCCCGAGAGGGGAAA 704  
 Db 708 ACTGCTTATGATGATCAACCCCGAGAGGGGAAA 741

RESULT 8  
 BF344334 697 bp mRNA linear EST 22-NOV-2000  
 LOCUS 602014708F1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4150512  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BF344334  
 VERSION BF344334.1 GI:11291554  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 697)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: David N. Louis, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9413 row: p column: 01  
 High quality sequence stop: 692.

FEATURES  
source  
Location/Qualifiers  
1. 697  
/organism="Homo sapiens"  
/db.xref="taxon:9606"  
/clone="IMAGE:4150512"  
/clone\_lib="NCI\_CGAP\_Brn64"  
/lab\_type="fibroblastoma with EGFR amplification"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1; NotI;  
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.57 kb. constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 173 a 167 c 220 g 137 t  
ORIGIN

Query Match 83.7%; Score 674.4; DB 12; Length 697;  
Best Local Similarity 98.7%; Pred. No. 5.4e-109;  
Matches 689; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

13 GGGGCCAAAGTGAAGTCCAGCGGTCTGCGACGCTTGGGCCACGCGCGCCCTGGGA 72  
1 GGGGCCAAAGTGAAGTCCAGCGGTCTTCCAGCGCTTGGGCCACGCGCGCCCTGGGA 60  
73 CCAAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGTTGGCTGCTGCT 132  
61 GCAGAGGTGGAGCAACCCGTTACCTTAARATGAAGGCTGGGTTGGCTGCTGCT 120  
133 TCTGGGGGCGCTGCTGGGAACCGCTGGGCTGGAGAGCGAGATCTCCACTGTGGAGC 192  
121 TCTGGGGGCGCTGCTGGGAACCGCTGGGCTGGAGAGCGAGATCTCCACTGTGGAGC 180  
193 ATGCAAGGCTCTGTGTGATGAATGGAATGGCAATTTGCCAGTGGAGACCCCAAGAGAC 252  
181 ATGCAAGGCTCTGTGTGATGAATGGAATGGCAATTTGCCAGTGGAGACCCCAAGAGAC 240  
253 CATTCAGATGGGATCTTTCGGATCAATCCAGATGGCCAGCTAGTGGTGGAGTGGC 312  
241 CATTCAGATGGGATCTTTCGGATCAATCCAGATGGCCAGCTAGTGGTGGAGTGGC 300  
313 TTATGCGCGCTGAGAGGCCCACTCAGACAGATGCTGGAGAGATATGACCGGATGAA 372  
301 TTATGCGCGCTGAGAGGCCCACTCAGACAGATGCTGGAGAGATATGACCGGATGAA 360  
373 GGAGTATGGGGAACAGATGATCTTCCACCCATCGCAAGAACTAGTGTAGTGG 432  
361 GGAGTATGGGGAACAGATGATCTTCCACCCATCGCAAGAACTAGTGTAGTGG 420  
433 CGGGAATGGAGATCCAGTGAATGGAGCTAGCAAGGCAATCCCAATCGACTGATATTAG 492  
421 CGGGAATGGAGATCCAGTGAATGGAGCTAGCAAGGCAATCCCAATCGACTGATATTAG 480  
493 CGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGGAGGAATAGAGATGATCAATTA 552  
481 CGGCAACCTCAAGTTTGCCTGTGAGAGCATTTGGAGGAATAGAGATGATCAATTA 540  
553 ATTCTTTTCCGAGAGGCTGACATGTTAAAGACAACTTTTCAGTAAAGCAAGATCT 612  
541 ATTCTTTTCCGAGAGGCTGACATGTTAAAGACAACTTTTCAGTAAAGCAAGATCT 600  
613 TTGTGACCATGCGCTGACATATCGCATGATGAGCATGGAACCACTGGAGCAG-CCACAC 672  
601 TTGTGACCATGCGCTGACATATCGCATGATGAGCATGGAACCACTGGAGCAG-CCACAC 659  
673 TGGCTTGAATGATCAACCCCGAGAGGGGAAATGGTGG 710  
660 TGGCTTGAATGATCAACCCCGAGAGGGGAAATGGTGG 697

RESULT 9 887 bp mRNA linear EST 25-SEP-2001  
BI762759 BI762759  
LOCUS 603048045F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5188307 5',  
DEFINITION mRNA sequence.

ACCESSION BI762759  
VERSION BI762759.1 GI:15754325  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 887)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LNL11470 row: 1 column: 12  
High quality sequence stop: 775.

FEATURES  
source  
Location/Qualifiers  
1. 887  
/organism="Homo sapiens"  
/db.xref="taxon:9606"  
/clone="IMAGE:5188307"  
/clone\_lib="NIH\_MGC\_116"  
/lab\_host="DH10B"  
/note="Organ: pooled colon, kidney, stomach; Vector:  
pCMV-Sport6; Site: 1; NotI; Site: 2; EcoRV (destroyed); RNA  
source anonymous pool of 3 colons, age 26 yo male, 49 yo  
female, 71 yo male colon; 46 yo male kidney, and pool of 2  
stomachs, 62 yo male and 70 yo female. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.4 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH\_MGC Library."

BASE COUNT 235 a 228 c 252 g 172 t  
ORIGIN

Query Match 83.3%; Score 671.4; DB 13; Length 887;  
Best Local Similarity 95.6%; Pred. No. 1.6e-108;  
Matches 743; Conservative 1; Mismatches 27; Indels 6; Gaps 5;

9 GGCTGGGGCCAAAGTGAAGTCCAGCGGTCTGCCAGCGCTGGGCCAGCGCGGCCCT 68  
60 GACTGGGGCCAAAGTGAAGTCCAGCGGTCTGCCAGCGCTGGGCCAGCGCGGCCCT 119  
69 GGAGCAAGGTGGAGCAACCCCGTTACCTTAARATGAAGGCTGGGTTGGCTGGCC 128  
120 GGGAGCAAGGTGGAGCAACCCCGTTACCTTAARATGAAGGCTGGGTTGGCTGGCC 179  
129 TGCCTCTGGGGGCCCTGTGCTGGAGACCGCTGGGCTGGAGAGCCAGGATCTCCACTGTG 188  
180 TGCCTCTGGGGGCCCTGTGCTGGAGACCGCTGGGCTGGAGAGCCAGGATCTCCACTGTG 239  
189 GAGCATGCAAGGCTGTGCTGATGAAGTGAATGGGAAATTTGCCAGGTGAACCCCAAGA 248  
240 GAGCATGCAAGGCTGTGCTGATGAAGTGAATGGGAAATTTGCCAGGTGAACCCCAAGA 299  
249 AAGCATTCAGATGGGATCTTCCGAGATCAATCCAGATGGAGCCAGTCAATGGTGGAGG 308  
300 AAGCATTCAGATGGGATCTTCCGAGATCAATCCAGATGGAGCCAGTCAATGGTGGAGG 359  
309 TGCCTTATGCGCGCTCAGAGGCCACCTCAAGAGCTGCTGGAGAGATATGTGACCGGA 368  
360 TGCCTTATGCGCGCTCAGAGGCCACCTCAAGAGCTGCTGGAGAGATATGTGACCGGA 419  
369 TGAAGAGATGGGAGACAGATGATCTTCCAGCCATCGCAAGAACTAGTACGTGTAG 428

Db 420 TGAAGAGTATGGGGAACAGATGATCTTCACCCATCGCAAGACTACCTAGCTTAG 479  
Qy 429 TGGGCCGAATGAGAAATCCAGTGAATGAGACTTACAAAGCATCCGAATCGACTAGATA 488  
Db 480 TGGGCCGAATGAGAAATCCAGTGAATGAGACTTACAAAGCATCCGAATCGACTAGATA 539  
Qy 489 TTAGGGGCAACCTCAAGTTTGGCGTGTGAGACATTTGTGGAGAAATACAGAGATGAACTCA 548  
Db 540 TTAGGGGCAACCTCAAGTTTGGCGTGTGAGAGCA-TGTGAGAGAAATACAGAGATGAACTCA 598  
Qy 549 TTGAATTTCTTCCCGAGAGCGTGCATGTTAAAGCAAACTTGCAGTAAGCAAG 608  
Db 599 TTGAATTTCTTCCCGAGAGCGTGCATGTTAAAGCAAACTTGCAGTAAGCAAG 658  
Qy 609 ATCTTTGTACCATGCGCTGCACATATGATGATGAGCTATGAAACCATGGAGAGCC 668  
Db 659 ATCTTTGTACCATGCGCTGCACATATGATGATGAGCTATGAAACCATGGAGAG-CC 717  
Qy 669 ACACGGCTTATGATGATACCCCGAGAGGGGAAATGG-TGGCAATGCTTTTATA-TA 726  
Db 718 ACACGGCTTATGATGATACCCCGAGAGGGGAAATGGTGGCAATGCTTTTATACTA 777  
Qy 727 TTATGTTTACT--GAATTAAGTAAATAATGAAACCAAAAGTAAAAA 781  
Db 778 TTATGTTTACTACTACATTAACCTGGGACACCTATGAAACCATTAAGTAAAAACA 834

RESULT 10  
B0668530 673 bp mRNA linear EST 15-JUL-2002  
LOCUS AGENCOURT\_8211118 NIH\_MGC\_102 Homo sapiens cdna clone IMAGE:6276413  
DEFINITION 5', mRNA sequence.  
ACCESSION B0668530  
VERSION B0668530.1 GI:21778777  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 673)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
plate: LCM2461 row: c column: 06  
High quality sequence start: 208  
High quality sequence stop: 350.  
location/Qualifiers  
1..673  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6276413"  
/clone\_lid="NIH\_MGC\_102"  
/tissue="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pOTB7; site\_1: XhoI;  
site\_2: EcoRI; cdna made by oligo-dt priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GCGACGAG(C) library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-cDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT  
ORIGIN

196 a 149 c 184 g 144 t

Query Match 83.1%, Score 669.8; DB 14; Length 673;  
Best Local Similarity 99.7%, Pred. No. 3.5e-108;  
Matches 671; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 118 TTGGCTGGCCCTGCTTCTGGGGGCCCTGCTGGGAACCCGCTGGGCTCGGAGAGCCAGCA 177  
Db 1 TTGGCTGGCCCTGCTTCTGGGGGCCCTGCTGGGAACCCGCTGGGCTCGGAGAGCCAGCA 60  
Qy 178 TCTCCACTGTGAGCATCAGAGGCTCTGTGGATGATAGAAATGGAAATGGCCAGCT 237  
Db 61 TCTCCACTGTGAGCATCAGAGGCTCTGTGGATGATAGAAATGGAAATGGCCAGCT 120  
Qy 238 GGACCCCAAGAAAGACATTCAGATGGATCTTCCGATCAATCCAGATGGCAGCCACTC 297  
Db 121 GGACCCCAAGAAAGACATTCAGATGGATCTTCCGATCAATCCAGATGGCAGCACTC 180  
Qy 298 AGTGTGAGAGTGGCTTATGCCCCCTCAGAGGCCCACTCAGAGCTGCTGGAGAGAT 357  
Db 181 AGTGTGAGAGTGGCTTATGCCCCCTCAGAGGCCCACTCAGAGCTGCTGGAGAGAT 240  
Qy 358 ATGTGACCGGATGAAAGAGTATGGGGAACAGATGATGATCTCCACCCATCGAAGACTA 417  
Db 241 ATGTGACCGGATGAAAGAGTATGGGGAACAGATGATGATCTCCACCCATCGAAGACTA 300  
Qy 418 CGTACGCTAGTGGGCGCGAATGGAATCCAGTGAACCTGACCTTCAAGGCTATCCGAAAT 477  
Db 301 CGTACGCTAGTGGGCGCGAATGGAATCCAGTGAACCTGACCTTCAAGGCTATCCGAAAT 360  
Qy 478 CGACTCAATATTTAGCCGACCCCTCAATTTGGCTGTAGAGCATTTGGAGAAATACGA 537  
Db 361 CGACTCAATATTTAGCCGACCCCTCAATTTGGCTGTAGAGCATTTGGAGAAATACGA 420  
Qy 538 GGATGAACCTCATGTAATCTTTTCCGAGAGCTGACATGTTAAAGCAAACTTTCAG 597  
Db 421 GGATGAACCTCATGTAATCTTTTCCGAGAGCTGACATGTTAAAGCAAACTTTCAG 480  
Qy 598 TTAGCGAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGATGATGAACAC 657  
Db 481 TTAGCGAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGATGATGAACAC 540  
Qy 658 TGGAGCACCCCACTGCTGTGATGATGATCACCACAGAGGGGAAATGTGGCAATGCC 717  
Db 541 TGGAGCACCCCACTGCTGTGATGATGATCACCACAGAGGGGAAATGTGGCAATGCC 600  
Qy 718 TTTTATATATATGTTTCTTACTGAAATTAAGTGAAGAAATGAAAGTAAAAA 777  
Db 601 TTTTATATATATGTTTCTTACTGAAATTAAGTGAAGAAATGAAAGTAAAAA 660  
Qy 778 AAAAAAAAAAAGA 790  
Db 661 AAAAAAAAAAAGA 673  
RESULT 11  
B1909759 720 bp mRNA linear EST 16-OCT-2001  
LOCUS B1909759  
DEFINITION 603070844F1 NIH\_MGC\_118 Homo sapiens cdna clone IMAGE:5219603 5',  
ACCESSION B1909759  
VERSION B1909759.1 GI:16173041  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 720)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LNL1552 row: a column: 12  
 High quality sequence stop: 717.

FEATURES  
 source location/Qualifiers  
 1..720

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5219603"  
 /clone\_1lb="NIH\_MGC\_118"  
 /tissue\_type="Leukocyte"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV  
 (destroyed); RNA source leukocytes from anonymous pool of  
 non-activated adult donors. Library is oligo-dT primed  
 and directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.7 kb, insert size range  
 1.2-3.3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 027. Note:  
 this is a NIH\_MGC Library."

BASE COUNT 176 a 175 c 230 g 139 t  
 ORIGIN

Query Match 82.0%; Score 660.8; DB 13; Length 720;  
 Best Local Similarity .97.8%; Pred. No. 1.3e-105;

Matches 679; Conservative 1; Mismatches 13; Indels 1; Gaps 1;

QY 11 CTGGGGCCAAAGTGAAGTCCAGGCGTGTGCGAGGCTTGGGCCCGCGGCGCCCTGG 70  
 DB 27 CTGGGGCCAAAGTGAAGTCCAGGCGTGTGCGAGGCTTGGGCCCGCGGCGCCCTGG 86  
 QY 71 GACCAAGGTGAGCAACCCGCTTACCTTAARATGAAGGCTGGGCTGGGCGCCCTG 130  
 DB 87 GAGCAGAGGTGAGCAGCAGCCCATTAAGTGAAGGCTGGGCTGGGCGCCCTG 146  
 QY 131 CTTGCGGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGAGCCAGATCTCCACTGTGA 190  
 DB 147 CTTCTGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGAGCCAGATCTCCACTGTGA 206  
 QY 191 GCATGCAAGGCTGCTGTGATGATGAATGGAATTTGCCAGGTGCAAGCCCAAGAG 250  
 DB 207 GCATGCAAGGCTGCTGTGATGATGAATGGAATTTGCCAGGTGCAAGCCCAAGAG 266  
 QY 251 ACCATTAGATGGATCTTTCCGGATCAATCCAGATGGCAGCAGTCACTGTGAGAGT 310  
 DB 267 ACCATTAGATGGATCTTTCCGGATCAATCCAGATGGCAGCAGTCACTGTGAGAGT 326  
 QY 311 CCTATATCCCGCTCAGAGGCGCCACCTCAGAGAGTGTGAGAGATATGTGACCGGATG 370  
 DB 327 CCTATATCCCGCTCAGAGGCGCCACCTCAGAGAGTGTGAGAGATATGTGACCGGATG 386  
 QY 371 AAGGAGTATGGGGAACAGATTGATCTTCCACCATCGAAGAACTCGTACGATGTG 430  
 DB 387 AAGGAGTATGGGGAACAGATTGATCTTCCACCATCGAAGAACTCGTACGATGTG 446  
 QY 431 GGCCGGAATGGAATTCATGTAATGAGACTTCAAGGATCCGAATCGACTGATAT 490  
 DB 447 GGCCGGAATGGAATTCATGTAATGAGACTTCAAGGATCCGAATCGACTGATAT 506  
 QY 491 AGCGGACCCCTCAAGTTGGCTGAGAGAGATTGGAGAAATACAGAGATGAATCAT 550  
 DB 507 AGCGGACCCCTCAAGTTGGCTGAGAGAGATTGGAGAAATACAGAGATGAATCAT 566  
 QY 551 GAATTTTTCCTCCGAGAGGCTGCAATGTTAAAGACAAATTTGCAAGTGAAGACAGAT 610  
 DB 567 GAATTTTTCCTCCGAGAGGCTGCAATGTTAAAGACAAATTTGCAAGTGAAGACAGAT 626  
 QY 611 CTTTGTGACCATGCGCTGCAC-ATATGCGATGATGAGTATGAAACCTGAGACAGCCA 669

DB 627 CTTGCGACCAATGCGCTGACACATATCGCATATGATGATGACCATGAGAGGCCA 686  
 QY 670 CACTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 703  
 DB 687 CACTGCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

RESULT 12  
 LOCUS BE799495 914 bp mRNA linear EST 20-SEP-2000  
 DEFINITION 601589289P1 NIH\_MGC\_7 Homo sapiens CDNA clone IMAGE:3943223 5',  
 mRNA sequence.  
 ACCESSION BE799495  
 VERSION BE799495.1 GI:10220693  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 914)  
 NIH\_MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
 Plate: LNC9798 row: b column: 24  
 High quality sequence start: 4  
 High quality sequence stop: 788.  
 Location/Qualifiers  
 1..914

FEATURES  
 source

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3943223"  
 /clone\_1lb="NIH\_MGC\_7"  
 /tissue\_type="small cell carcinoma"  
 /cell\_line="MGC3"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: Lung; Vector: pORF7; Site.1: XhoI; Site.2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 230 a 216 c 292 g 176 t  
 ORIGIN

Query Match 80.7%; Score 650.8; DB 12; Length 914;  
 Best Local Similarity 92.9%; Pred. No. 6.5e-105;  
 Matches 736; Conservative 1; Mismatches 48; Indels 7; Gaps 5;

QY 11 CTGGGGCCAAAGTGAAGTCCAGGCGTGTGCGAGGCTTGGGCCCGCGGCGCCCTGG 70  
 DB 127 CTGGGGCCAAAGTGAAGTCCAGGCGTGTGCGAGGCTTGGGCCCGCGGCGCCCTGG 186  
 QY 71 GACCAAGGTGAGCAACCCGCTTACCTTAARATGAAGGCTGGGCTGGGCGCCCTG 130  
 DB 187 GAGCAGAGGTGAGCAGCAGCCCATTAAGATGAAGGCTGGGCTGGGCGCCCTG 246  
 QY 131 CTTGCGGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGAGCCAGATCTCCACTGTGA 190  
 DB 247 CTTGCGGGGGGCGCTGCTGGGGAACCGGCTGGGCTGGGAGAGCCAGATCTCCACTGTGA 306  
 QY 191 GCATCAGGCGCTGCTGTGATGATGAATGGAATTTGCCAGTGTGAGACCCCAAGAG 250



Db	307	GCATCAGAGGCTCTGTGGATGTAAGTACAGAAATGGAAAAATTTGCCAGGTGGACCCCAAGAG	366
Oy	251	ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGGCATCAAGTGGTAGG	310
Db	367	ACCATTCAGATGGGATCTTTCCGGATCAATCCAGATGGCAGGCATCAAGTGGTAGG	426
Oy	311	CCWTATGCCCGCTCAGAGGCCACCTCAACAGAGCTGCTGGAGAGATATATGACCGGATG	370
Db	427	CCTTATGCCCGCTCAGAGGCCACCTCAACAGAGCTGCTGGAGAGATATATGACCGGATG	486
Oy	371	AAGGATATGGGGAAACAGATTGATCTTTCACACCATTCGCAMAGACTACGTATAGT	430
Db	487	AAGGATATGGGGAAACAGATTGATCTTTCACACCATTCGCAMAGACTACGTATAGT	546
Oy	431	GGCCGGATGGGAATCCAGTGAAGTGAACCTTCAGAAAGCATCCGAATGACATCAATAT	490
Db	547	GGCCGGATGGGAATCCAGTGAAGTGAACCTTCAGAAAGCATCCGAATGACATCAATAT	606
Oy	491	AGCGCACCCCTCAAGTTTGCCTGAGAGACATTTGGAGAGAAATACGAGATGAACATAT	550
Db	607	AGCGCACCCCTCAAGTTTGCCTGAGAGACATTTGGAGAGAAATACGAGATGAACATAT	666
Oy	551	G-AATTCCTTTCGAGAGGCTGACAAATGTTAAAGACAACTTTCAGATGAACGAGACA	609
Db	667	GAAATTCCTTTCGAGAGGCTGACAAATGTTAAAGACAACTTTCAGATGAACGAGACA	725
Oy	610	TCTTTGTACCATGCGCTGCACATATCCGATATAGCATATACCACTGGAGCAGCCCA	669
Db	726	TCTTTGTACCATGCGCTGCACATATCCGATATAGCATATACCACTGGAGCAGCCCA	784
Oy	670	CACATGCTTGATGATCAACCCCGAGAGGGGAAAAATGGTGCATATGCCCTTATATATTA	729
Db	785	CACATGCTTGATGAT- AACCCAGAGAGGGGAAAAATGGTGCATATATATATATAT- 842	
Oy	730	TGTTTACTGAATTAAGTAAAAAATATGAAACCAAAAGTAAAAAATAAAAAATAAAAAAG	789
Db	843	--GGTTACTGAATTAAGTAAAAAATGAAACCAAGTAAAAAATAAAAAATCGGGGATCG	900
Oy	790	AGAGAGAGAG 801	
Db	901	TGGGGGAGAG 912	

FEATURES	source
RESULT 13	
LOCUS	791 bp mRNA
DEFINITION	linear
ACCESSION	EST 07-MAY-2001
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

BG704443 791 bp mRNA  
 602688633f1 NIH\_MGC\_95 Homo sapiens CDNA clone IMAGE:4821008 5',  
 mRNA sequence.  
 BG704443  
 BG704443.1 GI:13977790  
 EST.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 791)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 plate: LLAM10727 row: a column: 09  
 High quality sequence stop: 785.  
 Location/Qualifiers  
 1..791

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:4821008"
/clone_id="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptpr (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (ctcgag
); Oligo-dn primed using primer 5'-tttcttttttttttttvtvn-3',
size selected for average insert size 2.5 kb and
normalized for KOF 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

```

Query Match	Score	DB	Length
Best Local Similarity	97.1%	Pred. No. 4,1e-104;	
Matches	668;	Conservative	1; Mismatches 17; Indels 2; Gaps 1;

  

QY	11	CTGGGCGCCAAAGTGAATCCACGCGGTGTGCCAGCGCTTGGGCGCCAGCGCGCGCCCTCG	70
DB	106	CTGGGCGCCAAAGTGAATCCACGCGGTGTGCCAGCGCTTGGGCGCCAGCGCGCGCCCTCG	165
QY	71	GACCAAGAGTGGAGAACCCCGTTACCCCTAAARATGAAGAGCTGGGGTGTGGCTGGCCCTG	130
DB	166	GAGCAGAGGTGGAGCGACGCCCTTATACGCTTAAGATGAAGAGCTGGGGTGTGGCTGGCCCTG	225
QY	131	CTTTCGGGGGGCCCTGCTGGGAACCGGCTCGGAGCGCCAGAGATCTCCACTGTGGA	190
DB	226	CTTTCGGGGGGCCCTGCTGGGAACCGGCTCGGAGCGCCAGAGATCTCCACTGTGGA	285
QY	131	GCATCAGAGGCTCTGTGTGGATGAACTAATAATGGGAAATGCCAGGTGGACCCCAAGAG	250
DB	286	GCATCAGAGGCTCTGTGTGGATGAACTAATAATGGGAAATGCCAGGTGGACCCCAAGAG	345
QY	251	ACCATTCAGATGGGATCTTTCGCGATCAATCCAGATGGCAGCCAGTCAGTGTGGAGGTG	310
DB	346	ACCATTCAGATGGGATCTTTCGCGATCAATCCAGATGGCAGCCAGTCAGTGTGGAGGTG	405
QY	311	CGTTATGCGCCGCTCAGAGGCGCACCTCAGAGAGCTGCTGGAAGAGATATGTGACCGATG	370
DB	406	CGTTATGCGCCGCTCAGAGGCGCACCTCAGAGAGCTGCTGGAAGAGATATGTGACCGATG	465
QY	371	AAGGAGTATGGGGGAACAATATGATCTTCCACACCATCCCAAGAACTAGTACGTATAG	430
DB	466	AAGGAGTATGGGGGAACAATATGATCTTCCACACCATCCCAAGAACTAGTACGTATAG	525
QY	431	GCGCGAGTATGGAGAAATCAGTGAACCTGGAACCTACAAAGGCTCCGAAATCAGATAT	490
DB	526	GCGCGAGTATGGAGAAATCAGTGAACCTGGAACCTACAAAGGCTCCGAAATCAGATAT	585
QY	491	AGCGGACCCCTCAAGTTTGCSTGTGAGAGCAATGTGAGAGAAATACGAGATGAATCTCAT	550
DB	586	AGCGGACCCCTCAAGTTTGCSTGTGAGAGCAATGTGAGAGAAATACGAGATGAATCTCAT	645
QY	551	GAATCTCTTTTCCCGAGAGGCTGACAAATTTAAAGCAAACTTTGCACTATACGAAACAGAT	610
DB	646	GAATCTCTTTTCCCGAGAGGCTGACAAATTTAAAGCAAACTTTGCACTATACGAAACAGAT	705
QY	611	CTTTGTGACAAATGCGCTCACATATCGCATATGAGTATGAGTATGAGTATGAGTATGAGTAT	670
DB	706	CTTTGTGACAAATGCGCTCACATATCGCATATGAGTATGAGTATGAGTATGAGTATGAGTAT	763
QY	671	ACTGCTTGATGATCAACCCCGAGAGG	698
DB	764	CATGTTTGATGATCAACCCCGAGAGG	791

  

RESULT	14
RF312281	



LOCUS BF312281 765 bp mRNA linear EST 21-NOV-2000  
DEFINITION 601898804F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4128049 5',  
mRNA sequence.  
ACCESSION BF312281  
VERSION BF312281.1 GI:11260080  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 765)  
AUTHORS NIH-MGC http://mgi.mcl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: L1CM1020 row: h column: 02  
High quality sequence stop: 724.  
Location/Qualifiers  
1..765  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4128049"  
/clone\_1lb="NIH\_MGC\_19"  
/tissue\_type="neuroblastoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: brain; Vector: pORF7; Site:1: XhoI; Site:2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
BASE COUNT 217 a 157 c 233 g 158 t  
ORIGIN

Query Match 80.0%; Score 645; DB 12; Length 765;  
Best Local Similarity 95.4%; Pred. No. 7.4e-104;  
Matches 707; Conservative 1; Mismatches 26; Indels 7; Gaps 4;

56 CGCGCGGCGCCCTGGACCAAGAGTGGAGCAACCCGCTTACCTTAARATGAAGGCTGG 115  
|||||  
2 CGCGCGGCGCCCTGGACGAGAGTGGAGCAACCCGCTTACCTTAAGATGAAGGCTGG 61  
|||||  
116 GATTGCTGGCCCTGCTTGGGGGCGCTGCTGGGAACGCGCTGGGCTGGAGAGCCAG 175  
|||||  
62 GATTGCTGGCCCTGCTTGGGGGCGCTGCTGGGAACGCGCTGGGCTGGAGAGCCAG 121  
|||||  
176 GATCTCCATCTGTGAGCATGCAAGGCTGTGTGATGAATAGAAATGGAATGGCCAG 235  
|||||  
122 GATCTCCATCTGTGAGCATGCAAGGCTGTGTGATGAATAGAAATGGAATGGCCAG 181  
|||||  
236 GTGGACCCCAAGAACATTCAGATGGGATCTTCCGGATCAATCAGATGGCAGCCAG 295  
|||||  
182 GTGGACCCCAAGAACATTCAGATGGGATCTTCCGGATCAATCAGATGGCAGCCAG 241  
|||||  
296 TCAGTGGTGGAGTGGCTTATGCGCCCTCAGAGAGCCACATCAGAGCTGTGGAGAG 355  
|||||  
242 TCAGTGGTGGAGTGGCTTATGCGCCCTCAGAGAGCCACATCAGAGCTGTGGAGAG 301  
|||||  
356 ATATGTGACCGGATGAAGAGATATGGGGAACAGATTGATCTTCCACCCATCGCAAGAC 415  
|||||  
302 ATATGTGACCGGATGAAGAGATATGGGGAACAGATTGATCTTCCACCCATCGCAAGAC 361  
|||||  
416 TACGTACGTGTATGTGGCGCGGATGGAATTCAGTGAATGGAATGGAATGGAATTCGA 475  
|||||

Db 362 TACGTACGTGTATGTGGCGCGGATGGAATTCAGTGAATGGAATGGAATGGAATTCGA 421  
Qy 476 ATGACTCAGATATATAGCGGACCCCTCAAGTTGGTGTGAGAGCATTTGGAGGAATAC 535  
|||||  
Db 422 ATGACTCAGATATATAGCGGACCCCTCAAGTTGGTGTGAGAGCATTTGGAGGAATAC 481  
Qy 536 GAGATGAAGTCAATTAATCTTTTCCGAGAGCTGACAAATGTTAAAGCAAACTTTCC 595  
|||||  
Db 482 GAGATGAAGTCAATTAATCTTTTCCGAGAGCTGACAAATGTTAAAGCAAACTTTCC 541  
Qy 596 AGTACGGAACAGATCTTTGTGACATGCCCTGACATATGCGATGATGATGATGATGAC 655  
|||||  
Db 542 AGTACGGAACAGATCTTTGTGACATGCCCTGACATATGCGATGATGATGATGATGAC 599  
Qy 656 ACTGAGAGCCACACATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 715  
|||||  
Db 600 ACTGAGAGCCACACATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 655  
Qy 716 CTTTAT 775  
|||||  
Db 656 GCTTAT 714  
Qy 776 AAAAAAAAAAAGAGAGAGA 796  
|||||  
Db 715 AAAAAAAAAAAGAGAGAGA 735

RESULT 15  
BE791763 745 bp mRNA linear EST 20-SEP-2000  
LOCUS 60181586F1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:3936141 5',  
mRNA sequence.  
DEFINITION BE791763  
ACCESSION BE791763  
VERSION BE791763.1 GI:10212961  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 745)  
AUTHORS NIH-MGC http://mgi.mcl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: DCTP/DRP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: L1CM779 row: k column: 22  
High quality sequence stop: 739.  
Location/Qualifiers  
1..745  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3936141"  
/clone\_1lb="NIH\_MGC-7"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pORF7; Site:1: XhoI; Site:2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 182 a 180 c 237 g 146 t  
ORIGIN

Query Match 79.58; Score 641; DB 12; Length 745;  
Best Local Similarity 97.08; Pred. No. 3.8e-103;  
Matches 684; Conservative 1; Mismatches 16; Indels 4; Gaps 3;

```
OY 11 CTGGGCGCAAGTGAAGTCCAGCGCTGCGCGGCTTGGGCGACGGCGCGCCCTGG 70
   |||||||
Db 45 CTGGGCGCAAGTGAAGTCCAGCGCTGCGCGGCTTGGGCGACGGCGCGCCCTGG 104
OY 71 GACCAAGGTGAGAGAACCCCGTTACCTAAARATGAAGGCTGGGGTTGGCTGGCCCTG 130
   |||||||
Db 105 GAGCAGAGGTGGAGCGACCCCATTAACGCTAAAGATGAAGGCTGGGGTTGGCTGGCCCTG 164
OY 131 CTTCTGGGGGGCCCTGCTGGGAAACCCCGTGGGCTGCGAGAGAGCCAGATCTCCACTGTGGA 190
   |||||||
Db 165 CTTCTGGGGGGCCCTGCTGGGAAACCCCGTGGGCTGCGAGAGAGCCAGATCTCCACTGTGGA 224
OY 191 GCATCAGAGGCTCTGTGGATGAATCTAGAAATGGGAAATTTGCCAGGTGACCCCAAGAAG 250
   |||||||
Db 225 GCATCAGAGGCTCTGTGGATGAATCTAGAAATGGGAAATTTGCCAGGTGACCCCAAGAAG 284
OY 251 ACCATTCAAGTGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAGTGTGGAGGTG 310
   |||||||
Db 285 ACCATTCAAGTGGATCTTTCCGGATCAATCCAGATGGCAGCCAGTCAGTGTGGAGGTG 344
OY 311 CCTTATGCGCGCTGAGAGGCCCACTCAAGAGCTGCTGGAGAGATATGTACCGGATG 370
   |||||||
Db 345 CCTTATGCGCGCTGAGAGGCCCACTCAAGAGCTGCTGGAGAGATATGTACCGGATG 404
OY 371 AAGGATATGGGGAACAGATTGATCTTCCACCCATCCGAAGACTAGTACGTGTAGTG 430
   |||||||
Db 405 AAGGATATGGGGAACAGATTGATCTTCCACCCATCCGAAGACTAGTACGTGTAGTG 464
OY 431 GGCCGGAATGGAGATCCAGTGAACCTGACCAAGGCATCCGAATCGACTCAGATATT 490
   |||||||
Db 465 GGCCGGAATGGAGATCCAGTGAACCTGACCAAGGCATCCGAATCGACTCAGATATT 524
OY 491 AGCGGCACCCCAAGTTGGCGTGAGAGCAATTTGGAGAGAAATAGAGATGAATCACTATT 550
   |||||||
Db 525 AGCGGCACCCCAAGTTGGCGTGAGAGCAATTTGGAGAGAAATAGAGATGAATCACTATT 584
OY 551 GAATTCCTTTCCCGAGAGGCTGACAAATGTAAAGACAACCTTTCAGTAAGCGAACAGAT 610
   |||||||
Db 585 GAATTCCTTTCCCGAGAGGCTGACAAATGTAAAGACAACCTTTCAGTAAGCGAACAGAT 644
OY 611 CTTTGTGACCATGCGCTGCACATATCGCATGTAGTGAACCATGAGACAGCCAC 670
   |||||||
Db 645 CTTTGTGACCATGCGCTGCACATATCGCATGTAGTGAACCATGAGACAGCC--A 701
OY 671 ACTGGCTGATGAGTCAACCCCGAGGAGGGAATAATGTGGCAATG 715
   |||||||
Db 702 CACTGGTTGATGATCA-CCCAAGAGGGGAATAATGTGGGCAGG 745
```

Search completed: March 30, 2003, 20:56:44  
Job time : 1155.74 secs

1:	/SID2/gcgdata/genseq/genseqn-emb1/NA1980.DAT.*
2:	/SID2/gcgdata/genseq/genseqn-emb1/NA1981.DAT.*
3:	/SID2/gcgdata/genseq/genseqn-emb1/NA1982.DAT.*
4:	/SID2/gcgdata/genseq/genseqn-emb1/NA1983.DAT.*
5:	/SID2/gcgdata/genseq/genseqn-emb1/NA1984.DAT.*
6:	/SID2/gcgdata/genseq/genseqn-emb1/NA1985.DAT.*
7:	/SID2/gcgdata/genseq/genseqn-emb1/NA1986.DAT.*
8:	/SID2/gcgdata/genseq/genseqn-emb1/NA1987.DAT.*
9:	/SID2/gcgdata/genseq/genseqn-emb1/NA1988.DAT.*
10:	/SID2/gcgdata/genseq/genseqn-emb1/NA1989.DAT.*
11:	/SID2/gcgdata/genseq/genseqn-emb1/NA1990.DAT.*
12:	/SID2/gcgdata/genseq/genseqn-emb1/NA1991.DAT.*
13:	/SID2/gcgdata/genseq/genseqn-emb1/NA1992.DAT.*
14:	/SID2/gcgdata/genseq/genseqn-emb1/NA1993.DAT.*
15:	/SID2/gcgdata/genseq/genseqn-emb1/NA1994.DAT.*
16:	/SID2/gcgdata/genseq/genseqn-emb1/NA1995.DAT.*
17:	/SID2/gcgdata/genseq/genseqn-emb1/NA1996.DAT.*
18:	/SID2/gcgdata/genseq/genseqn-emb1/NA1997.DAT.*
19:	/SID2/gcgdata/genseq/genseqn-emb1/NA1998.DAT.*
20:	/SID2/gcgdata/genseq/genseqn-emb1/NA1999.DAT.*
21:	/SID2/gcgdata/genseq/genseqn-emb1/NA2000.DAT.*
22:	/SID2/gcgdata/genseq/genseqn-emb1/NA2001.DAT.*
23:	/SID2/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*
24:	/SID2/gcgdata/genseq/genseqn-emb1/NA2002.DAT.*

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB	ID	Description
1	960	100.0	546	21	AAZ38326	Human transmembran
2	960	100.0	806	21	AAZ06869	Human secretory pe
3	960	100.0	806	21	AAZ08293	Human Zs19 gene e
4	960	100.0	814	21	AAZ38327	Human transmembran
5	960	100.0	814	24	ABK52765	CDNA encoding tran
6	960	100.0	814	24	ABL41995	Nucleotide sequenc
7	960	100.0	814	24	ABK09772	Human ovarian tumo
8	960	100.0	1180	20	AAZ22112	Human secreted pro
9	960	100.0	1210	22	AAZ21330	Human CDNA sequen
10	953	99.3	714	20	AAZ97884	Human secreted pro
11	938	97.7	1069	20	AAZ06870	Human secretory pe
12	938	97.7	1069	21	AAZ08294	Mouse secretory pe
13	914	95.2	657	24	ABK36007	Mouse ortholog gen
14	911	94.9	832	20	AAZ97837	CDNA sequence #398
15	835.5	87.0	592	21	AAA43525	Human secreted pro
16	393	40.9	484	22	AAZ56528	Mouse secreted exp
17	393	40.9	558	22	AAZ18023	Human breast cancer
18	393	40.9	595	22	AAZ07882	Human breast cancer
19	393	40.9	1160	22	AAZ26689	Human breast cancer
20	369	38.4	649	20	AAZ97883	Human secreted pro
21	369	38.4	649	20	AAZ06688	Human secretory pe
22	369	38.4	649	21	AAZ08284	Human Zs19 gene.
23	369	38.4	718	20	AAZ10655	CDNA encoding a hu
24	366	38.1	792	20	AAZ97836	Human secreted pro
25	365	38.0	415	21	AAZ06871	Zs19 expressed seq
26	365	38.0	415	21	AAZ08285	Expressed Sequence
27	355	37.0	1085	20	AAZ21330	WO99101020 Seq. ID 3
28	316	32.9	2532	22	ABA07341	Human pancreatic c
29	316	32.9	2532	22	AAZ32768	Human genomic DNA
30	316	32.9	2724	22	ABA07340	Human pancreatic c
31	316	32.9	2724	22	AAZ32767	Human genomic DNA
32	266	27.7	564	24	ABK09691	Human ovarian tumo
33	212	22.1	816	23	ABL21470	Drosophila melanog
34	184	19.2	352	22	ABA09488	Human secreted pro
35	154	16.0	786	21	AAZ37667	Human secreted pro
36	147	15.3	750	23	ABL16485	Arabidopsis thalia
37	147	15.3	2750	23	ABL16482	Drosophila melanog
38	147	15.3	3274	23	ABL16482	Drosophila melanog
39	141	14.7	3110	23	ABL21470	Drosophila melanog
40	131.5	13.7	744	21	AAZ49678	Human myocardiun s
41	131.5	13.7	746	22	AAZ94104	Primer specific fa
42	131.5	13.7	747	22	AAZ90704	Human secretory pro
43	131.5	13.7	1419	22	AAZ93914	Human CDNA encodin
44	131.5	13.7	1457	20	AAZ78923	Human zcalc-1 DNA
45	131.5	13.7	1462	21	AAA16621	Human secreted pro
ALIGNMENTS						
RESULT 1						
AAZ38326	ID	AAZ38326	standard; cDNA; 546 BP.			
XX	AAZ38326;					
XX	AC					
XX	DT					
XX	09-FEB-2000	(first entry)				
DE	Human transmembrane protein cDNA clone HP10390 coding sequence.					
XX						
XX	HP10390;	transmembrane domain; stomach cancer cell; antibody;				
KW	assay reagent;	diagnostic marker; primer; probe; antisense; gene therapy;				
KW	agonist; antagonist;	ligand; therapeutic; ds.				
XX						
OS	Homo sapiens.					
XX	Key	Location/Qualifiers				
FT	CDS	1..540				

```
FT      /*tag- a
FT      /product= "Human transmembrane protein HP10390"
XX      /note= "No stop codon given in the specification"
PN      W09955862-A2.
XX
XX      04-NOV-1999.
XX
XX      27-APR-1999; 99WO-JP02226.
XX
XX      28-APR-1998; 98JP-0119395.
XX
XX      (SAGA ) SAGAMI CHEM RES CENT.
XX      (PROT-) PROTEGENE INC.
XX
XX      Kato S, Kimura T;
XX
XX      WPI: 2000-023358/02.
XX      P-PSDB: AAY52391.
XX
XX      Human proteins with transmembrane domains, involved in control of cell
XX      proliferation and differentiation, useful for treating e.g. cancer or
XX      inflammation
XX
XX      Claim 3; Page 88; 114pp; English.
XX
XX      This sequence represents the coding sequence of human cDNA clone
XX      HP10390 which encodes a 20 kD protein with one putative transmembrane
XX      domain in the N-terminus. The cDNA was isolated from a human stomach
XX      cancer cell line cDNA library. The protein has no homology with any
XX      known protein. The protein may be used to raise specific antibodies, as
XX      assay reagents, as diagnostic tissue markers, for the isolation of
XX      cognate receptors, ligands and binding proteins, and as biologically
XX      active agents. Nucleotides encoding the protein may be used as primers
XX      and probes or antisense molecules, and in gene therapy. Cells transformed
XX      with these nucleotides may be used to screen for agonists and antagonists
XX      which are potentially useful therapeutically.
XX
XX      SO Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;

Alignment Scores:
Pred. NO.: 1.48e-103 Length: 546
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-082-502-17 (1-182) x AA238326 (1-546)
OY      1 MetlyvsglyttrpglyttrpleuAlaleuleucllyAlaleuleuGlyThrAlatrpaLa 20
DB      1 ATGAAAGGCTGGGGTGGCTGGCTGGCTGGGGGCCCTGCTGGGAAACCGCGCTGCT 60
OY      21 ArgatgserGlnAspLeuHnISCySgLYAlaCySArGAlaLeuValAspGluLeuclutrp 40
DB      61 CGGAGAGCCGAGATCTCCACTGTGGACATGCAAGGCTGTGGTGAAGATGAATGG 120
OY      41 GluilealaglnValAspProlyslsThrllleglmetGlyserPhearglleasnPro 60
DB      121 GAAATTTGCCAGTGGAGCCCAAGAGACATTCAGATGGATCTTTCGGATCAATGCA 180
OY      61 AspelyserGlnserValalGluValProfyralaArgserGluAlaHnIsleuThrglu 80
DB      181 GATGGACCCGCTCAGTGGTGGAGTGGCTTATGGCCCTCAGAGGCCACCTCACAGAG 240
OY      81 LeuengugluuileCySAspArgMetlysglyttrfyrglyuglnlleasproserThr 100
DB      241 CTGCTGGAGGAGATGTGACCGGATGAAGAGATGGGGAAACGATTGATCTTCCACC 300
OY      101 HlsarGlyAsanTyfValaJcValaValaGlyrGAnsglyluseSerGluIdeuAspLeu 120
DB      301 CATGCAAGAACTACATGATCTGTAGTGGCCGCGGAATGAGAAATCCAGTAACTGGACCTA 360
```

```
OY      121 GlnGlyllearGlleasPserAspIleSerGlyThrleuLyPheAlaCySgluSerlle 140
DB      361 CAAGCATCCGAAATCGACTCAGATATAGCGGACCCCTCAAGTTGCGTGGAGGACAT 420
OY      141 ValGluGlyttrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVallys 160
DB      421 GTGGAGAAATACGAGAGAACTCATTAATCTTTTCCGAGAGGCTGAGCAATCTTAAA 480
OY      161 AspLyLeuCySerLyArgThrAspLeuCySAspHlsAlaLeuHnIsleSerHlsAsp 180
DB      481 GACAAACTTTCAGTAAACGAAACAGATCTTTGTGACCATGCCCTGCACATATCCGATAT 540
OY      181 GluLeu 182
DB      541 GAGCTA 546

RESULT 2
ID      AAX06969 standard; cDNA; 806 BP.
XX
XX      AAX06969;
XX
XX      10-MAY-1999 (first entry)
XX
XX      Human secretory peptide-9 (Zsig9) variant cDNA.
XX
XX      Secretory peptide-9; Zsig9; human; tumour marker; cancer; therapy;
XX      diagnosis; growth enhancer; variant; ss.
XX
XX      Homo sapiens.
XX
XX      Key Location/Qualifiers
XX      CDS 104..652
XX      /tag= a
XX      sig_peptide 104..163
XX      /tag= b
XX      mat_peptide 164..649
XX      /tag= c

W09901554-A1.
PI      14-JAN-1999.
XX
XX      02-JUL-1998; 98WO-US13859.
XX
XX      17-JUN-1998; 98US-0099005.
XX      03-JUL-1997; 97US-0051704.
XX      03-JUL-1997; 97US-0088088.
XX      19-MAY-1998; 98US-0081338.
XX      19-MAY-1998; 98US-0085963.
XX      17-JUN-1998; 98US-0089899.
XX
XX      (ZYMO ) ZYMOGENETICS INC.
XX
XX      Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;
XX      WPI: 1999-106055/09.
XX      P-PSDB: AAW88474.
XX
XX      New mammalian secretory peptide-9 (Zsig9) - used as a growth
XX      enhancer for placenta, liver and heart, and as an indicator of
XX      cancer
XX
XX      Claim 2; Page 73-74; 85pp; English.
XX
XX      This cDNA clone encodes human secretory peptide-9, or Zsig9,
XX      variant (see AAW88474). Zsig9 (see also AAW88469) is overexpressed in
XX      human brain, liver, lung, oesophageal, stomach, colon, rectal,
XX      thyroid and lymphoma tumors. Thus, Zsig9 can be used as an
XX      indicator for cancer. Zsig9 cDNA was discovered in a placenta
XX      clone from a full-term pregnancy cDNA library which contained an
XX      expressed sequence tag (see AAX06971). The invention provides
```

CC polynucleotides (see AA06968-70) encoding Zs19 polypeptides (see  
CC AA08869-77) including mature polypeptides, other processed forms,  
CC variants and mouse orthologues. The Zs19 gene, or probes derived  
CC from it, can be used to determine if Zs19 is present on chromosome  
CC 12, and if a mutation has occurred. Antibodies raised against  
CC Zs19 can be used as diagnostic agents to determine the presence of  
CC Zs19, and thus the presence of cancer. They can also be labelled  
CC with radioisotopes or fused with toxins and used to treat tumours  
CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
CC proteins can be used to enhance the growth or development of the  
CC placenta, heart or liver.  
XX

SO Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;

## Alignment Scores:

Pred. No.:	2,49e-103	Length:	806
Score:	960.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-10-082-502-17 (1-182) x AA06969 (1-806)

QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTPala 20  
DB 104 ATGAAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 163  
QY 21 ArgArgSerGlnAspLeuHisGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 164 CGAGAGAGCCAGATCTCCCTCTGAGCATGCAAGGCTCTGGTGGTGAAGTAAAGTGG 223  
QY 41 GluIleAlaGlnValAspProLysTrpIleGlnMetGlySerPheArgIleAspPro 60  
DB 224 GAATATGCGCCAGCTGGAGCCCAAGAACATTCAGATGGATCTTCGGATCAATCCA 283  
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 284 GATGGAGCCAGTCAGTGGGAGGTGCTTATGCCCCGCTCAAGGCCACCTCACAGAG 343  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 344 CTCCTGAGAGAGATATGTACCGGATGAAGAGATATGGGAGACATTTGATCTTCACC 403  
QY 101 HisArgLysAsnTyrValArgValGlyArgAsnGlyLysSerGluLeuAspLeu 120  
DB 404 CATCGCAAGAACTACGTAGTGTAGTGGCCGGAATGGAATCCAGTGAAGTGAACCTA 463  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 464 CAAAGCATCCGAATCGACTCAGTATATACCGCACCCCTCAAGTTTCCGTGAGACATT 523  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 524 GTGAGAGATAGAGAGATGATGATCAATCTTTTCCGAGAGCTGACATGTTAA 583  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
DB 584 GACAAACTTGTGAGTAAAGCAAGATCTTGTGACCATGCCCTGCATATGCAATGAT 643  
QY 181 GluLeu 182  
DB 644 GACCTA 649

## RESULT 3

AAZ08293  
ID AAZ08293 standard; DNA; 806 BP.

AC AAZ08293;

DT 07-FEB-2000 (first entry)

XX

DE Human Zs19 gene encoding secretory protein variant-4.

XX Secretory protein-9; Human Zs19; chromosome 12q15 region; variant;  
KW overexpression; antagonist; antibody; antisense nucleotide; tumour;  
KW treatment; receptor; radio-label; fusion; polypeptide toxin; technique;  
KW down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;  
KW detection; stomach; lymphoma; alternative splicing; allelic variation;  
XX silent mutation; ds.

OS Homo sapiens.  
OS Synthetic.

XX Key Location/Qualifiers  
FH CDS 104..652

FT /tag= "a" Zs19 secretory protein variant-4"  
FT /note= "Overexpressed in tumours"

FT sig\_peptide 104..163  
FT /tag= b  
FT mat\_peptide 164..649  
FT /tag= c  
FT /label= Mature\_Zs19\_protein\_variant-4

XX W09960405-A1.

XX 25-NOV-1999.

XX 19-MAY-1999; 99WO-US11107.

XX 19-MAY-1999; 98DS-0081183.

XX (ZYMO ) ZYMOGENETICS INC.

XX Moore BE, Taft DW;

XX WPI. 2000-039447/03.

XX P-PSDB: AAY15135.

XX Detecting tumors using antibodies, antagonists and antisense  
XX nucleotides to secretory protein-9 (Zs19)

XX Disclosure: Page 33-35; 45pp: English.

XX The present DNA sequence is a gene encoding the variant of the secretory  
XX protein-9, Zs19 that arises due to alternative splicing, allelic  
XX variation or silent mutations that result in amino acid changes. This  
XX sequence is mapped to the human chromosome 12q15 region. It is  
XX overexpressed in tumours. Antagonists, antibodies and antisense  
XX nucleotides to Zs19 are useful for detecting and treating tumours. The  
XX antagonist may be an antibody or receptor to Zs19 and it may be radio-  
XX labelled or fused to a polypeptide toxin. It can be used for down  
XX regulating the overexpression of Zs19. The gene sequence can be used as  
XX a facilitates improved diagnostic and therapeutic techniques for detecting  
XX and treating cancers, especially of the brain, liver, stomach, lymphoma,  
XX etc., at an early stage.

SO Sequence 806 BP; 229 A; 183 C; 233 G; 160 T; 1 other;

## Alignment Scores:

Pred. No.:	2,49e-103	Length:	806
Score:	960.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-10-082-502-17 (1-182) x AAZ08293 (1-806)

QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTPala 20  
DB 104 ATGAAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 163

OY 21 ATGAGSerglnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 164 CGAGAGAGCCAGATCTCCACTGTGAGCATGACAGGGCTCTGTGATGAACTAAGATGG 223  
OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 224 GAAATTTGCCAGGTGGACCCCAAGAACATTCAGATGGATCTTTCCGGATCAATCCA 283  
OY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 284 GATGACAGCCAGTCACTGCTGGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 343  
OY 81 LeuLeuGluGlnIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 344 CTGCTGGAGAGATATGTACCGCATGAAAGAGATGGGGAACAGATTATCTTCCACCC 403  
OY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 404 CATCCCAAGAACTAGCTAGCTAGCTAGTGGCCGGAATGGGAATCCAGTGAACCTGACCTA 463  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 464 CAAGGCATCCGAAATGACATGATATAGCCGACCCCTCAAGTTGGCTGTGAGAGCATT 523  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 524 GTGGAGGAATACGAGATGAATCACTGATTAATCTTTTCCGAGAGGCTCAATGTGTTAA 583  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
DB 584 GACAAACTTTGAGTAAGCGAACAATCTTTGTGACATGCCCTGCACATATCGCATGAT 643  
OY 181 GluLeu 182  
DB 644 GAGCTA 649  
RESULT 4  
AAZ38327  
ID AAZ38327 standard; cDNA: 814 BP.  
AC AAZ38327;  
XX  
DT 09-FEB-2000 (first entry)  
DE Human transmembrane protein cDNA clone HP10390.  
XX  
KW HP10390; transmembrane domain; stomach cancer cell; antibody;  
KW assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;  
KW agonist; antagonist; ligand; therapeutic; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 145..693  
FT /tag="a  
FT /product="Human transmembrane protein HP10390"  
XX  
PN WO955862-A2.  
XX  
PD 04-NOV-1999.  
XX  
PF 27-APR-1999; 99WO-JP02226.  
XX  
PR 28-APR-1998; 98JP-0119395.  
XX  
PA (SAGA ) SAGAMI CHEM RES CENT.  
XX  
PI (PROT-) PROTEGENE INC.  
XX  
PI Kato S, Kimura T;  
XX  
DR WPI: 2000-023358/02.  
DR P-PSDB; AAY52391.  
XX

PT Human proteins with transmembrane domains, involved in control of cell  
PT proliferation and differentiation, useful for treating e.g. cancer or  
PT inflammation  
XX  
XX  
PS Claim 4; Page 106-107; 114pp; English.  
XX  
CC This sequence represents the human cDNA clone HP10390  
CC which encodes a 20 kD protein with one putative transmembrane  
CC domain in the N-terminus. The cDNA was isolated from a human stomach  
CC cancer cell line cDNA library. The protein has no homology with any  
CC known protein. The protein may be used to raise specific antibodies, as  
CC assay reagents, as diagnostic tissue markers, for the isolation of  
CC cognate receptors, ligands and binding proteins, and as biologically  
CC active agents. Nucleotides encoding the protein may be used as primers  
CC and probes or antisense molecules, and in gene therapy. Cells transformed  
CC with these nucleotides may be used to screen for agonists and antagonists  
CC which are potentially useful therapeutically.  
SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
Alignment Scores:  
Pred. No.: 2,53e-103 Length: 814  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 21  
US-10-082-502-17 (1-182) x AAZ38327 (1-814)  
OY 1 MetLysGlyTyrPglYThrLeuAlaLeuLeuGlyAlaLeuGlyThrAlaTrpAla 20  
DB 145 ATGAAAGCGTGGGTGGCTGCGCTCTGCGGGGCCCTGCGGAACCGCTGGGCT 204  
OY 21 ArgArgSerglnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 205 CGAGAGAGCCAGATCTCCACTGTGAGCATGACAGGCTCTGTGATGAACTAAGATGG 264  
OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 265 GAAATTTGCCAGGTGGACCCCAAGAACATTCAGATGGATCTTCCGGATCAATCCA 324  
OY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 325 GATGACAGCCAGTCACTGCTGGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 384  
OY 81 LeuLeuGluGlnIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 385 CTGCTGGAGAGATATGTGACCGCATGAAAGAGATGGGGAACAATGATCTTCCACC 444  
OY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 445 CATCCCAAGAACTAGCTAGCTAGTGGCCGGAATGGGAATCCAGTGAACCTGACCTA 504  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 505 CAAGGCATCCGAAATGACATGATTTGCGGACCCCTCAAGTTTGCCTGTGAGAGCATT 564  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 565 GTGGAGGAATACGAGATGAATCACTTGAATCTTTTCCGAGAGCTGACAAATGTGTTAA 624  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
DB 625 GACAAACTTTGAGTAAGCGAACAATCTTTGTGACATGCCCTGCACATATCGCATGAT 684  
OY 181 GluLeu 182  
DB 685 GAGCTA 690  
RESULT 5  
ABK52765  
ID ABK52765 standard; cDNA: 814 BP.

XX AC ABK52765;  
XX DT 15-AUG-2002 (first entry)  
XX DE cDNA encoding transmembrane protein 4, a cancer-linked protein.  
XX KM Expressed sequence tag; EST; human; cancer; anti-neoplastic;  
XX KW cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;  
XX transmembrane protein 4; gene; ss.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT CDS 145..693  
XX FT /\*tag= a  
XX FT /product= "Transmembrane protein 4"  
XX PN W0200231198-A2.  
XX PD 18-APR-2002.  
XX PE 11-OCT-2001; 2001WO-US31607.  
XX PR 11-OCT-2000; 2000US-239294P.  
XX PR 11-OCT-2000; 2000US-239297P.  
XX PR 11-OCT-2000; 2000US-239605P.  
XX PR 12-OCT-2000; 2000US-239802P.  
XX PR 12-OCT-2000; 2000US-239805P.  
XX PR 12-OCT-2000; 2000US-239806P.  
XX PR 16-OCT-2000; 2000US-240522P.  
XX PR 19-OCT-2000; 2000US-241682P.  
XX PR 19-OCT-2000; 2000US-241723P.  
XX PR 31-OCT-2000; 2000US-244932P.  
XX PA (AVAL-) AVALON PHARM.  
XX PI Young PE, Horriqan S, Weaver Z, Endress GA;  
XX DR WPI: 2002-463271/49.  
XX DR P-PSDB; AAU97063.  
XX PT Identifying modulators of a cancer-related gene to screen agents for  
XX PT preventing or treating cancer comprises detecting a difference in the  
XX PT expression of cancer-linked genes in the presence or absence of test  
XX PT compounds -  
XX PS Claim 1: Page 43; 66pp; English.  
XX CC The invention relates to modulators of a cancer-related genes. Also  
XX CC described are: (1) processes for identifying an anti-neoplastic agents  
XX CC comprising contacting a cell exhibiting neoplastic activity with a  
XX CC compound first identified as a cancer related gene modulator, and  
XX CC detecting a decrease in the neoplastic activity, (2) a process for  
XX CC determining the cancerous state of a cell by determining an increase in  
XX CC the level of expression of at least one gene, where an elevated  
XX CC expression relative to a known non-cancerous cell indicates a cancerous  
XX CC state or potentially cancerous state. The anti-neoplastic agent is  
XX CC useful for treating cancer or for protecting an animal against cancer.  
XX CC The immunogenic composition is also useful for treating cancer in an  
XX CC animal, where the composition elicits the production of cytotoxic T  
XX CC lymphocytes specific for the immunogenic composition. Preferably, the  
XX CC animal is a human. The cancer-linked genes and polypeptides are also  
XX CC useful as targets for cancer therapy or chemotherapy. The present  
XX CC sequence represents a cancer-linked gene located on chromosome 12,  
XX CC which encodes transmembrane protein 4.  
XX SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores: 2.53e-103 Length: 814  
Pred. No.: 960.00 Matches: 182  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00%

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
US-10-082-502-17 (1-182) x ABK52765 (1-814)  
QY 1 MetLysGlyTTrpGlyTTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
DB 145 ATGAAAGGCTGGGGGTGGCTGGCCCTGCTTGGGGGGCCCTGGGAAACCGCTGGCT 204  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 205 CGAGAGAGCCAGGATCTCCAGCTGGAGCATGCGAGGCTCTGGTGATAGACTAGAAATG 264  
QY 41 GluLeuAlaGlnValAspProLysTrpIleGlnMetGlySerPheArgIleAsnPro 60  
DB 265 GAAATGGCCAGGTGGAGCCCAAGAAAGACCATTCAGATGGATCTTCCGGATCAATCCA 324  
QY 61 AspGlySerGlnSerValValGluValProTyraAlaArgSerGluAlaHisLeuThrGlu 80  
DB 325 GATGCGAGCCAGTCAGTGGTGAGGTGCTTATGCCCGCTCAGAGGCCACCTCAGAGAG 384  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
DB 385 CTGCTGGAGAGATATATGTGACCGGATGAAGAGATATGGGAAACAGATTGATCTTCACCC 444  
QY 101 HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 445 CATCGCAGACACTACCTAGCTGTAGTGGCCGGAATGGAGAACTCCAGTAACTGGACCTTA 504  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 505 CAAGCATCCGAAATCGACATCAGATATATAGCGGACCCCTCAAGTTTCGCTGAGACATTT 564  
QY 141 ValGluGluTrpGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 565 GTGGAGGAATACAGAGATGAATCAATGAAATTTTCCCGAAGAGCTGCAATGTTAA 624  
QY 161 AspLysLeuCysSerLysArgTrpAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
DB 625 GACAACTTTGCGATAGAGCAACAGATCTTGTGACACATGCCCTGCACATATGCCATGAT 684  
QY 181 GluLeu 182  
DB 685 GAGCTA 690  
RESULT 6  
ABL41995  
ID ABL41995 standard; DNA; 814 BP.  
XX ABL41995;  
AC ABL41995;  
XX 11-JUN-2002 (first entry)  
DT 11-JUN-2002 (first entry)  
XX Nucleotide sequence of human polypeptide HP10390.  
DE Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.  
XX Homo sapiens.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 145..693  
FT /\*tag= a  
FT /product= "HP10390"  
XX W0200208416-A1.  
XX 31-JAN-2002.  
XX 24-JUL-2001; 2001WO-JP06371.  
XX 24-JUL-2000; 2000JP-0222743.  
XX 24-AUG-2000; 2000JP-0254407.





OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
 DB 145 ATGAAAGGCTGGGTTGGCTGGCCCTGCTTGTGGGGGCCCTGCTGGAAACCGCTGGCT 204  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlyTrp 40  
 DB 205 CGAGAGACCGAAGATCTCAGCTGTGACATGACAGGCGCTTGTGGATGAACCTGAAGTG 264  
 OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 265 GAAATGGCCAGGTGAGCCCAAGAAACATTCAGATGGATCTTCCGATCAATCA 324  
 OY 61 AspGlySerGlnSerValAlaGluValProGlyAlaAspSerGluAlaHisLeuThrGlu 80  
 DB 325 GATGGACACCGCTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCACAGAG 384  
 OY 81 LeuLeuGluGluIleCysAspArgMetLysGlyTrpGlyGluGlnIleAspProSerThr 100  
 DB 385 CTGCTGGAGGAGATATGTGACCGGATGAAGAGATGGGAACAGATTGATCTTCACAC 444  
 OY 101 HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 445 CATGGCAAGAACTACGACGTGTAGTGGCCGGAATGGAGAAATCCAGTGAACCTGACCTA 504  
 OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerTle 140  
 DB 505 CAAAGGATCCGAACTGCTCAGATGATTAAGGCGCACCTCAAGTTGGGTGTGAGACGATT 564  
 OY 141 ValGluGluTrpGlyLysAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB 565 GTGGAGGAATACGAGATGAACCTCATTTCTTTCCGAGAGCGCTGACATGTTTAA 624  
 OY 161 AspLysLeuLysSerLysValArgThrAspLeuLysAspHisAlaLeuHisIleSerHisAsp 180  
 DB 625 GACAAACTTGCAGTAAAGCAACAGATCTTGTGACCATGCTCCCTGACATATGCGATGAT 684  
 OY 181 GluLeu 182  
 DB 685 GAGCTA 690  
 RESULT 8  
 AAX22112  
 ID AAX22112 standard; DNA; 1180 BP.  
 AC AAX22112;  
 XX  
 DT 18-MAY-1999. (first entry)  
 DE Human secreted protein gene 2 clone H2MB56.  
 XX  
 KW Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
 KW tumour; neurodegenerative disorder; leukemia; autoimmune disease; AIDS;  
 KW developmental abnormality; foetal deficiency; Alzheimer's disease;  
 KW cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
 KW immune deficiency disease; respiratory disorder; arthritis; skeletal;  
 KW haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
 KW cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
 OS Homo sapiens.  
 XX  
 PN WO9901020-A2.  
 XX  
 PD 14-JAN-1999.  
 XX  
 PF 30-JUN-1998; 98MO-US13608.  
 XX  
 PR 12-SEP-1997; 97US-0058663.  
 PR 01-JUL-1997; 97US-0051381.  
 PR 01-JUL-1997; 97US-0051480.  
 PR 12-SEP-1997; 97US-0058598.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;  
 XX WPI: 1999-105683/09.  
 DR P-PSDB: AAY01136, AAY01162, AAY01163.  
 DR  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, immune deficiency diseases or blood  
 PS disorders  
 PS Claim 4; Page 117; 179pp; English.  
 XX  
 XX The invention relates to nucleic acid sequences (AAX22111 to AAX22134)  
 CC encoding human secreted proteins (AAY01135 to AAY01158). The secreted  
 CC protein gene sequences are deposited with the ATCC under deposit number  
 CC ATCC 209118. Host cells comprising recombinant vectors containing the  
 CC nucleic acid sequences are used for the recombinant production of the  
 CC secreted proteins. The polynucleotide and amino acid sequences are useful  
 CC for are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Pathological conditions can  
 CC be also diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC CC In, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorder, immune deficiency diseases  
 CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC haematopoietic disorders, neural disorders, cardiovascular disorders,  
 CC osteoporosis or gastrointestinal disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The present sequence represents a  
 CC gene encoding a human secreted protein (see descriptor line for gene  
 CC number and clone identification).  
 XX  
 SQ Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other:  
 Alignment Scores:  
 Pred. No.: 4, 16e-103 Length: 1180  
 Score: 960.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0  
 US-10-082-502-17 (1-182) x AAX22112 (1-1180)  
 OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
 DB 513 ATGAAAGGCTGGGTTGGCTGGCCCTGCTTGTGGGGGCCCTGCTGGAAACCGCTGGCT 572  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlyTrp 40  
 DB 573 CGAGAGACCGAAGATCTCAGCTGTGACATGACAGGCGCTTGTGGATGAACCTGAAGTG 632  
 OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 633 GAAATGGCCAGGTGAGCCCAAGAAACATTCAGATGGATCTTCCGATCAATCA 692  
 OY 61 AspGlySerGlnSerValAlaGluValProGlyAlaAspSerGluAlaHisLeuThrGlu 80  
 DB 693 GATGGACACCGCTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCACAGAG 752  
 OY 81 LeuLeuGluGluIleCysAspArgMetLysGlyTrpGlyGluGlnIleAspProSerThr 100  
 DB 753 CTGCTGGAGGAGATATGTGACCGGATGAAGAGATGGGGAACAGATTGATCTTCACAC 812  
 OY 101 HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 813 CATGGCAAGAACTACGACGTGTAGTGGCCGGAATGGAGAAATCCAGTGAACCTGACCTA 872  
 OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerTle 140

```

Db      873  CAAGGCATCCGATGACATGATATAGCGCCACCCCTCAACTTTGCGGTGAGACAT 932
Oy      141  ValGluGluTYrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVallys 160
Db      993  GTGGAGGAATACGAGAGATGAACTCATGTATCTTTCCCGAGAGCGTACAAATCTTAAA 992
Oy      161  AspLysLeuCYSerLysArgThrAspLeuCYAspHisAlaLeuHisIleSerHisasp 180
Db      993  GACAAACTTTCAGTAGACGAGACAGATCTTGTGACCATGCGCTGCACATATCGCATGAT 1052
Oy      181  GluLeu 182
Db      1053  GAGCTA 1058

RESULT 9
AAS21330
ID  AAS21330 standard; cDNA: 1210 BP.
XX
XX  AAS21330;
XX
XX  24-OCT-2001 (first entry)
XX
XX  Human cDNA sequence encoding for PRO4426 polypeptide.
DE
XX
XX  Human secretory and transmembrane; PRO; mammalian; cancer; lung;
KW  breast; prostate; cervical; tumour necrosis factor- $\alpha$ ; TNF- $\alpha$ ;
KW  cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
KW  adipocyte; A-peptide; factor VIIa; gene therapy; ss.
XX
XX  Homo sapiens.
OS
XX  WO200140466-A2.
XX
XX  07-JUN-2001.
XX
XX  01-DEC-2000; 2000WO-US32678.
XX
XX  01-DEC-1999; 99WO-US28301.
XX
XX  01-DEC-1999; 99WO-US28634.
XX
XX  02-DEC-1999; 99WO-US28551.
XX
XX  02-DEC-1999; 99WO-US28564.
XX
XX  02-DEC-1999; 99WO-US28565.
XX
XX  09-DEC-1999; 99US-0170262.
XX
XX  16-DEC-1999; 99WO-US30095.
XX
XX  20-DEC-1999; 99WO-US30911.
XX
XX  20-DEC-1999; 99WO-US30999.
XX
XX  30-DEC-1999; 99WO-US31243.
XX
XX  06-JAN-2000; 2000WO-US00277.
XX
XX  06-JAN-2000; 2000WO-US00376.
XX
XX  11-FEB-2000; 2000WO-US03565.
XX
XX  18-FEB-2000; 2000WO-US04341.
XX
XX  18-FEB-2000; 2000WO-US04342.
XX
XX  22-FEB-2000; 2000WO-US04914.
XX
XX  24-FEB-2000; 2000WO-US04914.
XX
XX  24-FEB-2000; 2000WO-US05004.
XX
XX  01-MAR-2000; 2000WO-US05601.
XX
XX  20-MAR-2000; 2000WO-US07377.
XX
XX  21-MAR-2000; 2000WO-US07532.
XX
XX  30-MAR-2000; 2000WO-US08439.
XX
XX  17-MAY-2000; 2000WO-US13705.
XX
XX  22-MAY-2000; 2000WO-US14042.
XX
XX  30-MAY-2000; 2000WO-US14941.
XX
XX  02-JUN-2000; 2000WO-US15264.
XX
XX  10-NOV-2000; 2000WO-US30873.
XX
XX
XX  (GETH ) GENENTECH INC.
XX
XX  Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
XX  Gerlisen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S,
XX  Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX  WPI; 2001-408281/43.

```

```

DR      P-PSDB; AAU12258.
XX
XX  Isolated, secretory and transmembrane PRO polypeptide used to detect
XX  other PRO polypeptides, link bioactive molecules to cells expressing
XX  PRO polypeptides, and detect the presence of mammalian tumours e.g.
XX  lung, breast, prostate, cervical
XX
XX  Claim 3; Fig 173; 813pp; English.
XX
XX  AAS21244-AAS21518 encode for novel human secretory and transmembrane
XX  PRO polypeptides. The PRO polypeptides are useful to detect other
XX  PRO polypeptides, to link bioactive molecules to cells expressing
XX  PRO polypeptides, to modulate biological activities of cells expressing
XX  PRO polypeptides, and to detect the presence of mammalian lung, colon,
XX  breast, prostate, rectal, cervical or liver tumours by comparing PRO
XX  polypeptide expression in a cell sample to that in a control sample.
XX  Some of the 275 sequences are also useful to stimulate the release of
XX  tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) from human blood, the
XX  proliferation or differentiation of chondrocytes, the proliferation or
XX  gene expression in pericyte cells, the release of proteoglycans from
XX  cartilage, the proliferation of inner ear utricular supporting cells or
XX  of T-lymphocytes, the release of a cytokine from peripheral blood
XX  monocytes (PBMCs), or the proliferation of endothelial cells. Some of
XX  the PRO polypeptides may modulate glucose or free fatty acid uptake by
XX  skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
XX  to factor VIIa. The PRO polypeptides can be used in assays to identify
XX  molecules involved in binding interactions. The polynucleotides encoding
XX  PRO polypeptides can be used to generate probes, antisense RNA/DNA,
XX  transgenic or knock out animals and can be used in gene therapy.
XX
XX  Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;
XX
XX  Alignment Scores:
XX  Pred. No.: 4,31e-103 Length: 1210
XX  Score: 960.00 Matches: 182
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 100.00% Indels: 0
XX  DB: 22 Gaps: 0
XX
XX  US-10-082-502-17 (1-182) x AAS21330 (1-1210)
Oy      1  MetLysGlyTTPGlyTTPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20
Db      543  ATGAAGGCTGGGTGGCTGGCCCTGCTTGGGGGCCCTGCGAACCCTGGGGCT 602
Oy      21  ArgArgSerGlnAspLeuHisCYSerGlyValaCYAspArgAlaLeuValaGluLeuGluTrp 40
Db      603  CGGAGGACCCAGATCTCCTCAGTGTGAGACATGACAGGCTTGGTGATGAATCTAGAAATGG 662
Oy      41  GluIleAlaGluValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db      663  GAAATTTGCCAGGTGAGACCCCAAGAAACCATTTGAGATGGGATCTTCCGGATCAATCCA 722
Oy      61  AspGlySerGlnSerValaValaGluValaProGlyAlaArgSerGluAlaHisLeuThrGlu 80
Db      723  GATGCGACCGCATCTGCTGTGAGGTGCTTATGCGCTTACGAGGCCACCTCACAGAG 782
Oy      81  LeuLeuGluGluIleCYAspArgMetLysGlyTTPGlyGluGluGlnIleAspProSerThr 100
Db      783  CTGCTGGAGGAGATATGTGACCGGATGAGAGATGTGGGAAACATGATGATCTTCCACC 842
Oy      101  HisArgLysAsnTYrValaArgValaValaGluYrArgAsnGlyLysSerGluLeuAspLeu 120
Db      843  CATGCAAGAACTACGTACGTGTAGTGGCCGGAATGAGATCAGTAATCTGAGACTA 902
Oy      121  GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCYGluSerIle 140
Db      903  CAAGGCATCCGAATCGATCAGATATATAGCGCCACCCCTCAAGTTTGGCTGTGAGAGCAT 962
Oy      141  ValGluGluTYrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnVallys 160
Db      963  GTGGAGGAATACGAGAGATGAACTCATGTATCTTTCCCGAGAGCGTACAAATCTTAAA 1022

```

QY 161 Asplysleucylserlysarqthraspleucysasphialaleuhtisleserhisasp 180  
 DB 1023 GACAAACTTGGACGTAAAGCAGATCTTGTGACCATGCCCTGCACATATCGCATGAT 1082  
 QY 181 Glutreu 182  
 DB 1083 GAGCTA 1088  
 RESULT 10  
 AAX97884 standard; cDNA; 714 BP.  
 AC AAX97884;  
 DT 23-SEP-1999 (first entry)  
 DE Human secreted protein encoding cDNA #72.  
 DE Human secreted protein encoding cDNA #72.  
 KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
 KM diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
 OS Homo sapiens.  
 PN W0925825-A2.  
 XX  
 PD 27-MAY-1999.  
 PF 13-NOV-1998; 98WO-IB01862.  
 XX  
 PR 04-SEP-1998; 98US-0099273.  
 PR 13-NOV-1997; 97US-0066677.  
 PR 17-DEC-1997; 97US-0069957.  
 PR 09-FEB-1998; 98US-0074121.  
 PR 13-APR-1998; 98US-0081563.  
 PR 10-AUG-1998; 98US-0096116.  
 XX  
 PA (GENE) GENSET.  
 XX  
 PI Bougueleret L, Duclert A, Dumas Mline Edwards J;  
 DR WPI; 1999-347472/29.  
 DR P-PSDB; AAY36200.  
 XX  
 PT Extended cDNAs encoding secreted proteins  
 XX  
 PS Claim 1; Page 272; 307pp; English.  
 XX  
 CC AAX97813-X97906 represent extended cDNA's which encode novel human  
 CC secreted proteins (see AAY36129-Y36222) and which have cytosolic,  
 CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
 CC express secreted proteins or parts of them or to obtain antibodies  
 CC capable of binding to the secreted proteins. They may also be used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC Uses also include design of expression vectors and secretion vectors.  
 XX  
 SQ Sequence 714 BP; 203 A; 158 C; 202 G; 151 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,416-102 Length: 714  
 Score: 953.00 Matches: 181  
 Percent Similarity: 99.45% Conservative: 0  
 Best Local Similarity: 99.45% Mismatches: 1  
 Query Match: 99.27% Indels: 0  
 DB: 20 Gaps: 0  
 US-10-082-502-17 (1-182) x AAX97884 (1-714)  
 QY 1 MetlysglytyrpglytyrpleualaleuLeuGlyAlaLeuLenglyThralatrrpAla 20  
 DB 33 ATGAAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 92  
 QY 21 ArgargserglinspleuhtiscysglyalacysargalaleuvalaspluLeuGluTrp 40

DB 93 CGAGAGAGCCAGATATCCAGCTGGAGCATGAGGAGCTGTGGATGACATGAAATG 152  
 QY 41 GluIIAlaIaIValAspprolysthrIleGlmetylerpheargileanpro 60  
 DB 153 GAAATGGCCAGGTGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCA 212  
 QY 61 AspglyserclinservAlaIguValProtyrAlaargserclualahisLeuhtgIu 80  
 DB 213 GATGGCAGCCAGTCAGTGTGGAGGTGCTTATGCCGCTCAGAGGCCACCTCAGAG 272  
 QY 81 LeuLeuGluGluIleCysasparqmetlysglytyrGlyGluGluIleaspproserThr 100  
 DB 273 CTCCTGAGAGAGATATGTGACCGAGTGAAGAGATATGGGAAACAGATTATCTTCCAC 332  
 QY 101 HisArglyAsnArgValArgValAlaGlyArganglyGluSerSerGluLeuAspleu 120  
 DB 333 CATCGCAGAACTACCTAGCTGTAGTGGCCCGGAAGGAATCCAGTGAATGGACCTA 392  
 QY 121 GlnGlyIleArgIleasparqmetlyserGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 393 CAAGCATCCGATCGATCGATCATGATATTAGCGCACCTCAAGTTTGCCTGTGGAGCAT 452  
 QY 141 ValGluGluTyrgluAspGluLeuIleGluPheSerArgGluAlaAspAsnVallys 160  
 DB 453 GTGGAGGAATACAGAGATGACATTCATTAATCTTTCCGAGAGGCTGACATGTAA 512  
 QY 161 Asplysleucylserlysarqthraspleucysasphialaleuhtisleserhisasp 180  
 DB 513 GACAAACTTGGACGTAAAGCAGATCTTGTGACCATGCCCTGCACATATCGCATGAT 572  
 QY 181 Glutreu 182  
 DB 573 GAGCTA 578  
 RESULT 11  
 AAX06970 standard; cDNA; 1069 BP.  
 ID AAX06970;  
 AC AAX06970;  
 XX  
 DT 10-MAY-1999 (first entry)  
 XX  
 DE Mouse secretory peptide-9 (Zs19) orthologue cDNA.  
 DE Mouse secretory peptide-9 (Zs19) orthologue cDNA.  
 KW Secretory peptide-9; Zs19; orthologue; mouse; tumour marker;  
 KW cancer; therapy; diagnosis; growth enhancer; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 358..906 /\*tag= a  
 FT sig\_peptide 358..417 /\*tag= b  
 FT mat\_peptide 418..903 /\*tag= c  
 FT  
 FT  
 W09901554-A1.  
 PN  
 PD 14-JAN-1999.  
 XX  
 PF 02-JUL-1998; 98WO-US13859.  
 XX  
 PR 17-JUN-1998; 98US-0099005.  
 PR 03-JUL-1997; 97US-0051704.  
 PR 03-JUL-1997; 97US-0888088.  
 PR 19-MAY-1998; 98US-0081338.  
 PR 19-MAY-1998; 98US-0085983.  
 PR 17-JUN-1998; 98US-0089899.  
 PA (ZYMO) ZYMOGENETICS INC.

PI Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;  
 XX WPI: 1999-106055/09.  
 DR P-PSDB; AAM88476.  
 XX  
 PT New mammalian secretory peptide-9 (Zs19) - used as a growth  
 PT enhancer for placenta, liver and heart, and as an indicator of  
 PT cancer  
 PS Claim 2; Page 75-77; 85pp; English.  
 XX  
 CC This cDNA clone encodes novel mouse secretory peptide-9, or Zs19  
 CC (see AAM88476), an orthologue of novel human Zs19 (see AAM88465).  
 CC Human Zs19 is overexpressed in a number of tumours including  
 CC brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid  
 CC and lymphoma tumors, and thus can be used as an indicator for  
 CC cancer. The invention provides polynucleotides (see AAM88465-70)  
 CC encoding Zs19 polypeptides (see AAM88469-77) including mature  
 CC polypeptides, other processed forms, variants and the mouse  
 CC orthologues. The Zs19 gene, or probes derived from it, can be  
 CC used to determine if Zs19 is present on chromosome 10, and if a  
 CC mutation has occurred. Antibodies raised against Zs19 can be  
 CC used as diagnostic agents to determine the presence of Zs19, and  
 CC thus the presence of cancer. They can also be labelled with  
 CC radioisotopes or fused with toxins and used to treat tumours  
 CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.  
 XX  
 SO Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 1,41e-100 Length: 1069  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.71% Indels: 0  
 DB: Gaps: 20  
 US-10-082-502-17 (1-182) x AAX06970 (1-1069)  
 OY 1 MetLysGIYTPRGLYTPRLeuAlaLeuLeuGlyAlaLeuLeuGlyThraAlaTPPa 20  
 DB 358 ATGAAGGCTGGGGTGGTGGTACCTACTTTGGGGGTCTCTGGCAATCGCGGCT 417  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuAlaValAspGluLeuGluTPR 40  
 DB 418 CGAAGGAGCGCAAGATCTACACTGTGAGCTTCAGAGGCTCTGTGATGAATTAAGTGG 477  
 OY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 478 GAAATTTGCCCGGTGGAGCCCAAGAACCAATTCAGATGGGATCCCTCCGATCAATCCA 537  
 OY 61 AspGlySerGlnSerValValGluValProtyrAlaArgSerGlnAlaHisLeuThrGlu 80  
 DB 538 GATGGCAGCCAGCTAGTTGGAGGTACTTATGCCCGCTCAGAGGCCACCTCAGAGAG 597  
 OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
 DB 598 TTGCTTGAGGAGGTGTGACGAGATGAAGAGTACGGGGAACAAATGACCTTCTTACC 657  
 OY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
 DB 658 CACCCCAAGAACTAGCTAGCGCTCTGAGCCGAAATGGAATTCAGAACTAGACTTA 717  
 OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 718 CAGGCAATCCGAATGATCAGATATCAGCGCACCTCAAGTTTCCGTGTGAGAGCATTT 777  
 OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB 778 GTGGAAGAAATCAGAGGATGAGCTTATCGAATTTCTTCCAGAGAGGCTGACAACTTAA 837

OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisIleSerHisAsp 180  
 DB 838 GACAAACTTTCAGATGAAGCGAGACAGATCTATGACCAATGCCCTGCACAGATCTACGAT 897  
 OY 181 GluLeu 182  
 DB 898 GAGCTA 903  
 RESULT 12  
 AAZ08294  
 ID AAZ08294 standard: DNA; 1069 BP.  
 AC AAZ08294;  
 XX  
 XX 07-FEB-2000 (first entry)  
 DE Mouse ortholog gene encoding Zs19 secretory protein.  
 XX  
 KW Secretory protein-9; mouse Zs19; ortholog; overexpression; antagonist;  
 KW antibody; antisense nucleotide; tumour; treatment; receptor;  
 KW radio-label; polypeptide toxin; down-regulation; diagnostic;  
 KW therapeutic; probe; cancer; brain; liver; detection; stomach;  
 KW lymphoma; ds.  
 XX  
 OS Mus musculus.  
 Key Location/Qualifiers  
 FH 358..906  
 FT CDS  
 FT  
 FT /\*tag= a  
 FT /product= "Mouse Zs19 secretory protein"  
 FT /note= "Overexpressed in tumours"  
 FT sig\_peptide 358..417  
 FT /\*tag= b  
 FT mat\_peptide 418..903  
 FT /\*tag= c  
 FT /label= Mature\_Zs19\_protein\_variant-4  
 PN W09960405-A1.  
 XX  
 XX 25-NOV-1999.  
 PD  
 PF 19-MAY-1999; 99WO-US11107.  
 XX  
 PR 19-MAY-1998; 98US-0081183.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Moore EE, Taft DW;  
 XX  
 DR WPI: 2000-039447/03.  
 DR P-PSDB; AAY15136.  
 XX  
 PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (Zs19)  
 PS  
 PS Disclosure; Page 37-38; 45pp; English.  
 XX  
 CC The present DNA sequence is an ortholog encoding the secretory protein-9,  
 CC Zs19 derived from mouse. It is overexpressed in tumours. Antagonists,  
 CC antibodies and antisense nucleotides to Zs19 are useful for detecting  
 CC and treating tumours. The antagonist may be an antibody or receptor to  
 CC Zs19 and it may be radio-labelled or fused to a polypeptide toxin. It  
 CC can be used for down regulating the overexpression of Zs19. The gene  
 CC sequence can be used as nucleic acid probes to detect RNA encoding Zs19.  
 CC The Zs19 sequence facilitates improved diagnostic and therapeutic  
 CC techniques for detecting and treating cancers, especially of the brain,  
 CC liver, stomach, lymphoma etc., at an early stage.  
 XX  
 SO Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 1,41e-100 Length: 1069

Score:	98.00	Matches:	177
Percent Similarity:	98.35%	Conservative:	2
Best Local Similarity:	97.25%	Mismatches:	3
Query Match:	97.71%	Indels:	
DB:	21	Gaps:	0

OY 145 GUAAPGJLUleuIleGluPheSerArgGluAlaAspAsnValLysAspLysLeuGys 164  
|||||  
DB 421 GAGATGAACATCATGATTTCTTCCGAGAGCTGACCAATGTTAAAGACAAACTTTGCC 480  
OY 165 SerLysArgThrAspLeuGysAspHisAlaLeuHisIleSerHisAspGluLeu 182  
|||||  
DB 461 AGTAAGCGAAGATCTTTGTGACCATGCGCCCTGCACATATCCGATGATGAGCTA 534  
RESULT 14  
ID AAX97837 standard; cDNA: 832 BP.  
AC AAX97837;  
XX 23-SEP-1999 (first entry)  
DE Human secreted protein encoding cDNA #25.  
XX  
KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
OS Homo sapiens.  
XX  
PN W09925825-A2.  
XX  
PD 27-MAY-1999.  
XX  
PF 13-NOV-1998; 98WO-1B01862.  
XX  
PR 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
PA (GEST ) GENSET.  
XX  
PI Bouqueleret L, Duclet A, Dumas Milne Edwards J;  
XX  
DR WPI: 1999-347472/29.  
DR P-PSDB: AAY36153.  
XX  
XX Extended cDNAs encoding secreted proteins  
XX  
PS Example 28; Page 194-195; 307pp; English.  
XX  
XX AAX97813-X97906 represent extended cDNA's which encode novel human  
CC secreted proteins (see AAY36129-136222) and which have cytosolic,  
CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
CC express secreted proteins or parts of them or to obtain antibodies  
CC capable of binding to the secreted proteins. They may also be used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC Uses also include design of expression vectors and secretion vectors.  
XX  
SQ Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other;  
Alignment Scores:  
Pred. NO.: 1.5e-97 Length: 832  
Score: 911.00 Matches: 180  
Percent Similarity: 97.84% Conservative: 1  
Best Local Similarity: 97.30% Mismatches: 1  
Query Match: 94.90% Indels: 3  
Gaps: 0  
US-10-082-502-17 (1-182) x AAX97837 (1-832)  
OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyAlaLeuGlyTrpAlaTrpAla 20  
|||||  
DB 148 ATGAAGAGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 207  
OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGlu-LeuGluTr 40

DB 208 CGAGAGAGCCAGCATCTCCACTGTGAGACATGACGGCTGTGGATGAAACTGAGATG 267  
OY 40 pGluIleAlaGlnValAspProLysTrpThrIleGlnMetGlySerPheArgIleAsp 60  
DB 268 GGAATATGCCAGGTGGACCCAGAGAGACCTTGTAGATGGATTTTCGGATCAATCC 327  
OY 60 cAspLysSerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrG 80  
DB 328 AGATGGAGCCAGTCAGTGGTGGAGTGCCTTATCCCGCTAGAGGCCCACTCACA 387  
OY 80 uLeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSer 100  
DB 388 GCTCTCGAGAGATATGTGACCGGATGAGAGATGAGGAGAACATGATTCCTTCAC 447  
OY 100 rHisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGlySerSerGluLeuAsp 120  
DB 448 CCATCGCAGAACTACGTACGTAGTGGCCGGAATGAGATCCACTGAACTGGACCT 507  
OY 120 uGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLys-PheAlaCysGlnSer 140  
DB 508 ACAAGGCAATCCGAATCGACTGATATTAGCGGACCCCTCAAGBTTCGTGTGGAA 567  
OY 140 lValAlaGluGluTrpGluAspGluLeuIleGluPheSerArgGluAlaAspAsnVal 160  
DB 568 TTGTGAGAGAAATACGAGATGACATGAAATCTTTCCGAGAGGCTGCAATGTGA 627  
OY 160 yAspLysLeuGysSerLysArgThrAspLeuGysAspHisAlaLeuHisIleSer 179  
DB 628 AAGACAACTTTGCACTGATGAGCAACAGATCTTTGTGACCAATGCCCTGCACATATG 687  
OY 180 AspGluLeu 182  
DB 688 GATGAGCTA 696  
RESULT 15  
ID AAA43525 standard; cDNA: 592 BP.  
AC AAA43525;  
XX  
XX 21-AUG-2000 (first entry)  
DE  
XX  
XX Mouse secreted expressed sequence tag SEQ ID NO:100.  
DE  
KW Human; mouse; chicken; rat; secreted expressed sequence tag; sBST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokinetin; analgesic; haemostatic;  
KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;  
KW antiviral; antidiabetic; antisthmatic; vulnerrary; antiparkinsonian;  
KW antifibrinolytic; osteoprotective; neuroprotective; nontropic; antiparasitic;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
XX  
XX Mus musculus.  
OS  
PN W0200021991-A1.  
PD 20-APR-2000.  
XX  
XX 15-OCT-1999; 99WO-US24206.  
XX  
XX 15-OCT-1998; 98US-0104436.  
XX  
XX (GEMT ) GENETICS INST INC.  
PA  
XX  
PI Jacobs K, McCoy JM, Lavallic ER, Collins-Racie LA, Evans C;



PI Merberg D, Treacy M, Bowman MR;  
XX WPI: 2000-317938/27.  
PT Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (SESTs), useful for treating various disorders  
XX such as autoimmune, infectious, and central nervous system disorders -  
PS Claim 1: Page 228-229; 803pp; English.  
XX  
CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat  
CC tissue sources. The SESTs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC cytostatic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
CC antiasthmatic; vulnerrary; antilucer; osteopathic; neuroprotective;  
CC nociceptive; antiparkinsonian; antipsoriatic; cerebroprotective;  
CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
CC therapy and in vaccines. The SESTs are useful as probes for the  
CC identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
XX in the exemplification of the present invention.  
SQ Sequence 592 BP; 159 A; 139 C; 167 G; 123 T; 4 other;  
Alignment Scores:  
Pred. No.: 7.2e-89 Length: 592  
Score: 835.50 Matches: 165  
Percent Similarity: 93.48% Conservative: 3  
Best Local Similarity: 91.85% Mismatches: 10  
Query Match: 87.03% Indels: 4  
DB: 21 Gaps: 0  
US-10-082-502-17 (1-182) x AAA43525 (1-592)  
QY 1 MetTysGlyTrpPglYTrpLeuAla-leuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAl 20  
DB 44 ATGAAAGAGNTGGGGTGGCTAGGCCCTACTTTGGGGGCTCTGCGAAGCTGCTGGGA 103  
QY 20 AATGATGSeRgInAsPleuHISCySgIyAlaCySArgAlaLeuValAsPgluLeuGluTr 40  
DB 104 TCGAAGAGGCCAANATCTACACTGTGAGANCTTCAGAGGCTCTGATGATTAAGATG 163  
QY 40 pGluIleAlaGlnValAsPProLysLysThrIleGlnMetGlySerPheArgIleAsnPr 60  
DB 164 GGAATTTGCCCGGTGGACCCCAAGAACATTCAGATGGGATCCTTCGAATCAATCC 223  
QY 60 oAsPglYSeRgInSeRValValGluValProLyraAlaArgSeRgIuAlaHISLeuThrGl 80  
DB 224 AGATGGCACCACACTGCTGTGGAGGTACCTATGCCCCGTCAGAGGCCACCTCACAGA 283  
QY 80 uLeuLeuGluGluIleCyAsPArgMetLysGluTrgIyGluGlnIleAsPProSeRTh 100  
DB 284 GTTGTCTTGGAGAGTGTGTGACCGAATGAAGAGTACGGGA-CAGATTGA-CCTTCTAC 341  
QY 100 rHISArgLysAsnTyValArgValValGlyArgAsnGlyGluSeRgIuLeuAsPle 120  
DB 342 CCACCGCAAGAACTACGTACGCGTGTGACCGGATGAGATCCAGTGAACCTAGACTT 401  
QY 120 uGlnGlyIleArgIleAsPSeRAsPleSeRgIyThrLeuLysPheAlaCySgluSerI 140  
DB 402 ACAGGGCATCCGAATTGACTCATCATCANCGCACCTCAAGTTTGCGTGTGAGAGCAT 461

QY 140 eValGluGluTrgIuAsPgluLeuIleGluPheSeRArgIuAlaAsPAsnVally 160  
DB 462 TGTGGAAATAACGAGGATGAGCTTATCGAATTTCTTCCAGAGAGGCTGCAACGTTAA 521  
QY 160 sAsPlySeuCySeRlySArgThrAsPleuCyAsPhtAlaLeuHISleSer-HISA 180  
DB 522 AGACAACTTTCAGTAAAGCGGACAGATCTATGTGACCATCCCTCGACAGATCTCAGC 581  
QY 180 sPgluLeu 182  
DB 582 AAGAGCTC 589  
Search completed: March 30, 2003, 22:42:53  
Job time: 145.849 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 39.6802 Seconds

(without alignments)  
1406.624 Million cell updates/sec

Title: US-10-082-502-17

Perfect score: 960  
Sequence: 1 MKGGMWALLGALIGTAMA.....LCNRKTDICDHALHSHDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xip  
-O=/cgn2\_1/USPRO.spool/US10082502/umatch\_24032003\_135100\_6616/app-query.fasta.1.1308  
-DB=Issued\_Patents\_NA -OEMT=fastap -SOFTX=ini -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCCALIGM=200 -THR\_SCORE=Pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10082502.ECGN.1.1.131 etumatch\_24032003\_135100\_6616 -NCPU=6 -ICPU=3  
-NO\_XIPY -NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCUTS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	131.5	13.7	744	US-09-163-285-3	Sequence 3, Appl1
2	131.5	13.7	1512	US-09-163-285-1	Sequence 1, Appl1
3	131.5	10.3	915	US-09-724-864-14	Sequence 14, Appl1
4	85.5	8.9	61663	US-09-453-702B-62	Sequence 62, Appl1
5	85	8.8	5181	US-08-257-073-10	Sequence 10, Appl1
6	84	8.8	1924	US-08-961-083-159	Sequence 159, Appl1
7	83	8.6	3752	US-08-661-527-208	Sequence 208, Appl1
8	81.5	8.5	1216	US-09-071-035-3	Sequence 3, Appl1
9	81.5	8.5	1347	US-09-071-035-1	Sequence 1, Appl1
10	81.5	8.5	4394	US-09-103-152A-1	Sequence 1, Appl1
11	81.5	8.5	4403765	US-09-103-840A-2	Sequence 2, Appl1
12	81.5	8.5	4411529	US-09-103-840A-1	Sequence 1, Appl1

13	81	8.4	833	2	US-08-837-029-1	Sequence 1, Appl1
14	79.5	8.3	1173	3 <td>US-08-993-380-3<td>Sequence 3, Appl1</td></td>	US-08-993-380-3 <td>Sequence 3, Appl1</td>	Sequence 3, Appl1
15	78.5	8.2	5661	4 <td>US-08-938-105-2<td>Sequence 2, Appl1</td></td>	US-08-938-105-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
16	78	8.1	1530	4 <td>US-09-131-831B-2<td>Sequence 2, Appl1</td></td>	US-09-131-831B-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
17	78	8.1	1554	3 <td>US-08-809-999D-1<td>Sequence 1, Appl1</td></td>	US-08-809-999D-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
18	78	8.1	1554	3 <td>US-09-069-637-1<td>Sequence 1, Appl1</td></td>	US-09-069-637-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
19	78	8.1	1554	4 <td>US-09-322-360-1<td>Sequence 1, Appl1</td></td>	US-09-322-360-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
20	78	8.1	1554	4 <td>US-09-131-831B-1<td>Sequence 1, Appl1</td></td>	US-09-131-831B-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
21	78	8.1	1555	2 <td>US-08-809-999D-2<td>Sequence 1, Appl1</td></td>	US-08-809-999D-2 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
22	78	8.1	2148	3 <td>US-09-069-637-2<td>Sequence 2, Appl1</td></td>	US-09-069-637-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
23	78	8.1	2148	3 <td>US-08-485-355B-47<td>Sequence 2, Appl1</td></td>	US-08-485-355B-47 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
24	78	8.1	2148	4 <td>US-09-322-360-2<td>Sequence 2, Appl1</td></td>	US-09-322-360-2 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
25	77.5	8.1	1776	2 <td>US-08-522-229B-1<td>Sequence 2, Appl1</td></td>	US-08-522-229B-1 <td>Sequence 2, Appl1</td>	Sequence 2, Appl1
26	77.5	8.1	1776	2 <td>US-09-102-433-1<td>Sequence 1, Appl1</td></td>	US-09-102-433-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
27	77.5	8.1	2508	4 <td>US-08-850-993-1<td>Sequence 1, Appl1</td></td>	US-08-850-993-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
28	77	8.0	3218	4 <td>US-09-221-017B-255<td>Sequence 255, Appl1</td></td>	US-09-221-017B-255 <td>Sequence 255, Appl1</td>	Sequence 255, Appl1
29	76.5	8.0	2012	2 <td>US-08-484-200-3<td>Sequence 3, Appl1</td></td>	US-08-484-200-3 <td>Sequence 3, Appl1</td>	Sequence 3, Appl1
30	76.5	8.0	2478	4 <td>US-08-485-355B-47<td>Sequence 4, Appl1</td></td>	US-08-485-355B-47 <td>Sequence 4, Appl1</td>	Sequence 4, Appl1
31	76.5	8.0	2478	4 <td>US-08-485-355B-49<td>Sequence 4, Appl1</td></td>	US-08-485-355B-49 <td>Sequence 4, Appl1</td>	Sequence 4, Appl1
32	76.5	8.0	2478	4 <td>US-09-194-613-1<td>Sequence 1, Appl1</td></td>	US-09-194-613-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
33	76.5	8.0	2479	4 <td>US-08-485-355B-51<td>Sequence 1, Appl1</td></td>	US-08-485-355B-51 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
34	76.5	8.0	4865	3 <td>US-08-894-017-24<td>Sequence 1, Appl1</td></td>	US-08-894-017-24 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
35	76	7.9	4131	1 <td>US-07-872-644-38<td>Sequence 38, Appl1</td></td>	US-07-872-644-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
36	76	7.9	4131	1 <td>US-08-297-494-38<td>Sequence 38, Appl1</td></td>	US-08-297-494-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
37	76	7.9	4131	1 <td>US-08-297-510-38<td>Sequence 38, Appl1</td></td>	US-08-297-510-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
38	76	7.9	4131	1 <td>US-08-479-532-38<td>Sequence 38, Appl1</td></td>	US-08-479-532-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
39	76	7.9	4131	1 <td>US-08-455-526-38<td>Sequence 38, Appl1</td></td>	US-08-455-526-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
40	76	7.9	4131	1 <td>US-08-455-526-38<td>Sequence 38, Appl1</td></td>	US-08-455-526-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
41	76	7.9	4131	1 <td>US-09-139-491-38<td>Sequence 38, Appl1</td></td>	US-09-139-491-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
42	76	7.9	4131	5 <td>PCT-US92-03222-38<td>Sequence 38, Appl1</td></td>	PCT-US92-03222-38 <td>Sequence 38, Appl1</td>	Sequence 38, Appl1
43	75	7.8	5558	4 <td>US-08-961-527-103<td>Sequence 103, Appl1</td></td>	US-08-961-527-103 <td>Sequence 103, Appl1</td>	Sequence 103, Appl1
44	75	7.8	6519	1 <td>US-08-588-985-1<td>Sequence 1, Appl1</td></td>	US-08-588-985-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1
45	75	7.8	6519	1 <td>US-08-971-988-1<td>Sequence 1, Appl1</td></td>	US-08-971-988-1 <td>Sequence 1, Appl1</td>	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-163-285-3  
Sequence 3, Application US/09163285  
Patent No. 6204013

GENERAL INFORMATION:  
APPLICANT: Khodadoust, Mehran  
TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHYE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,285  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 60/090,398  
FILING DATE: June 24, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragoras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MNI-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..744  
US-09-163-285-3

Alignment Scores:  
Pred. No.: 2,828-07 Length: 744  
Score: 131.50 Matches: 50  
Percent Similarity: 44.74% Conservative: 35  
Best Local Similarity: 26.32% Mismatches: 62  
Query Match: 13.70% Indels: 43  
DB: Gaps: 11

US-10-082-502-17 (1-182) x US-09-163-285-3 (1-744)

OY 7 LeuAlaLeuLeuGlyAlaLeuLeuGlyThr-----AlaTrpAla-----Arg 21  
||| :|||  
DB 16 TTGGGAATATTCCTTTCTTCTTTTGGCCGCGACAGAGCCTTGGCGATGTTGAAG 75  
OY 22 ArgSerGlnAsp-----LeuHisCysGlyAlaCysArgAlaLeuVal 35  
||| :|||  
DB 76 GAGGAGACGATGACACAGAACGCTTCCACCAATGCGAAGTGTGTAACTGCTGAGC 135  
OY 36 AapGluLeuGluTrpGluLeuLeuGluValAsp---ProLysLysThrLeuGlnMetGly 54  
||||| :|||  
DB 136 ACAGAGCTACAGAGCGGAACTGATGCGACCGGTCATCTCGAGAGTGTGGAGTGGGG 195  
OY 55 SerPheArgIleAsnProAspGlySerGlnSerValAlaGluValProTyrAlaArgSer 74  
||| :|||  
DB 196 CAGTGTGCTG-----GATACAGGCAAGAGAGAGACAGCTGCTTACAGCGTTTCA 246  
OY 75 GluAlaHisLeuThrGluLeuLeuGluIleCysAspArgMetLysGluTyrGlyGlu 94  
||| :|||  
DB 247 GAGACAGAGCTGGAAGAGCCCTTAGAATTTATGTGACCGGATCTCGACTAT----- 300  
OY 95 GlnIleAspProSerThrHis-----ArgLysAsnTyrValArgValAlaIlyArgAsn 112  
||| :|||  
DB 301 -----AGTGTTCACGCTGACGCGCAGGCTCAGTAGATATGACC-----AAG 342  
OY 113 GlyIleuSerSerGlu-----LeuAspLeuGlnGlyIleArgIleAsp 126  
||| :|||  
DB 343 GGTGAGAGTACAGACCATGGCAACACTGAAGGCTTAGTGCAAGAGGGGTGAAGTGGAT 402  
OY 127 SerAspIleSer-----GlyThrLeuLysPhe 135  
||| :|||  
DB 403 CTGGGAGATCCCTCTGAGCTTTGGGATGAGCCAGCGTGAGATCACTACCTCAAGAG 462  
OY 136 AlaCysGluSerIleValGluIleuTrpGluAspGluLeuIleGlu---PhePheSerArg 154  
||||| :|||  
DB 463 CAGTGTGAGACCATTTGGAGAGATTGAAAGCATTTGGGAGACTGGTACTTCCACCAT 522  
OY 155 GluAlaAspAsnValLysAspLysLeuGly 164  
||| :|||  
DB 523 CAGGAGCAGCCCTACAAATAATTTCTCTGT 552

RESULT 2  
US-09-163-285-1  
Sequence 1, Application US/09163285  
Patent No. 6204013  
GENERAL INFORMATION:  
APPLICANT: Rhodadoust, Mehran  
TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street

CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,285  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/090,398  
FILING DATE: June 24, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragoras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: KNI-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1512 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 125..868  
US-09-163-285-1

Alignment Scores:  
Pred. No.: 7,848-07 Length: 1512  
Score: 131.50 Matches: 50  
Percent Similarity: 44.74% Conservative: 35  
Best Local Similarity: 26.32% Mismatches: 62  
Query Match: 13.70% Indels: 43  
DB: Gaps: 11

US-10-082-502-17 (1-182) x US-09-163-285-1 (1-1512)

OY 7 LeuAlaLeuLeuGlyAlaLeuLeuGlyThr-----AlaTrpAla-----Arg 21  
||| :|||  
DB 140 TTGGGAATATTCCTTTCTTCTTTTGGCCGCGACAGAGCCTTGGCGATGTTGAAG 199  
OY 22 ArgSerGlnAsp-----LeuHisCysGlyAlaCysArgAlaLeuVal 35  
||| :|||  
DB 200 GAGGAGACGATGACACAGAACGCTTGGCCACCAATGCGAAGTGTGTAACTGCTGAGC 259  
OY 36 AapGluLeuGluTrpGluLeuLeuGluValAsp---ProLysLysThrLeuGlnMetGly 54  
||||| :|||  
DB 260 ACAGAGCTACAGAGCGGAACTGATGCGACCGGTCATCTCGAGAGTGTGGAGTGGGG 319  
OY 55 SerPheArgIleAsnProAspGlySerGlnSerValAlaGluValProTyrAlaArgSer 74  
||| :|||  
DB 320 CAGTGTGCTG-----GATACAGGCAAGAGAGAGACAGCTGCTTACAGCGTTTCA 370  
OY 75 GluAlaHisLeuThrGluLeuLeuGluIleCysAspArgMetLysGluTyrGlyGlu 94  
||||| :|||  
DB 371 GAGACAGAGCTGGAAGAGCCCTTAGAATTTATGTGACCGGATCTCGACTAT----- 424  
OY 95 GlnIleAspProSerThrHis-----ArgLysAsnTyrValArgValAlaIlyArgAsn 112  
||| :|||  
DB 425 -----AGTGTTCACGCTGAGCGCAAGGCTCAGTAGATATGACC-----AAG 466  
OY 113 GlyIleuSerSerGlu-----LeuAspLeuGlnGlyIleArgIleAsp 126  
||| :|||  
DB 467 GGTGAGTACAGACCATGGCAACACTGAAGGCTTAGTGCAAGAGGGGTGAAGTGGAT 526

Oy 127 Seraspleser-----GlyThleuylsPhe 135  
Db 537 CTGGGATCCCTCGACCTTTGGGATGAGCCACGCTGAGGTACATACCTCAAGAA 586  
Oy 136 AlacysgiserilevalglutyrGluaspGluileglu---PhePheSerarg 154  
Db 587 CAGGTGTAGACACATTTGGAGAGAGTTTAAACATTTGTGGAGAGCTGTACTTCCACCAT 646  
Oy 155 GluAlaspaspaValLysAspLysLeuCs 164  
Db 647 CAGAGAGAGCCCTCAACAAATTTCTCTCT 676  
RESULT 3  
US-09-724-864-14  
Sequence 14, Application US/09724864  
Patent No. 6380362  
GENERAL INFORMATION:  
APPLICANT: Watson, James D.  
APPLICANT: Murlson, James G.  
TITLE OF INVENTION: Polynucleotides, polypeptides expressed  
FILE REFERENCE: 11000.105001  
CURRENT APPLICATION NUMBER: US/09/724,864  
PRIOR FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 915  
TYPE: DNA  
ORGANISM: Mouse  
US-09-724-864-14  
Alignment Scores:  
Pred. No.: 0.00414 Length: 915  
Score: 99.00 Matches: 50  
Percent Similarity: 39.71% Conservative: 31  
Best Local Similarity: 24.51% Mismatches: 73  
Query Match: 10.31% Indels: 50  
Gaps: 10  
US-10-082-502-17 (1-182) x US-09-724-864-14 (1-915)  
Oy 7 LeuAlaleuLeuLeuGly-----AlaleuLeuGlyThrlaTrpAlaArgArgSer--- 23  
Db 38 CTGTGTCTACTGTCTGGGTGAGGCTATCTGGGAGCGCCGGGATAGGTTCCCTC 97  
Oy 24 -----GlnAsp 25  
Db 98 TCGGCTTCGGCTCCACACATGATGAAGAGAGTACTGGCTCATATGCCGGCTCAC 157  
Oy 26 LeuHicysGlyAlaLacysArgAlaLeuValaAspLyuLeuGluTrpGluIleAlaGluVal 45  
Db 158 CTCGGCTGCGATCTGCGGCTGCGCTTCCAGATGGGGAAGCTGTGGGAAAGCA 217  
Oy 46 AspProLysLysThrIleGlnMetGlySerPheArgIleasnProAspGlySerGlnSer 65  
Db 218 GAGGCTAAACCTGAC-----ACTCCAGACGCCAGT---GGA 250  
Oy 66 ValValGluValaProTyrAlaArgSerGluAlaHicLeuThrluLeuLeuGluGluIle 85  
Db 251 TTGACAGAGAGCTG-----AGTGAATCCACGATACCAATGTCCCGGACACAGAC 298  
Oy 86 CysAspArg---MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysasn 104  
Db 299 TGTCTTCAGAACTGGACGCTCTATGGA-----GTTCATGAAGTGAAC 340  
Oy 105 TyrValaArgValaGluArgaangLysLysSerGluLeuAspLeuGlnGlyIleArg 124  
Db 341 CAGATGAAGCTGTCAAGCGGCCAGACCTTAGCAAGGGCCAGAG-----CCAAAG 391  
Oy 125 IleAspSerAspLysSerGly-----ThleuLysPheAlaLacysGluSer 139

Db 392 ATCAGCGTGTATGATTTGGGGGTCCCTGGCCCAATAGAGCTCTCCAGACCTGTTCCAC 451  
Oy 140 IleValGluGluTyr---GluAspLyuLeuIleGluPhePheSerArgGluAlaAspasn 158  
Db 452 TACTGTGTAGTTGGAGAGACAGATCTATGAAGCTTACCCGCAAGGCCAAGCGMAT 511  
Oy 159 ValLysAspLysLeuCsSerLysAlaArgThrAspLeuCsAspPheAlaLeuHicIleSer 178  
Db 512 CTGAGAGGGCTGCTGTGTGGGGGACCAACCATGGGCCCTGCTCAGAGAGATCTGGCCAG 571  
Oy 179 HisaspLyuLeu 182  
Db 572 AGAGAGAGCTT 583  
RESULT 4  
US-09-453-702B-62  
Sequence 62, Application US/09453702B  
Patent No. 6365723  
GENERAL INFORMATION:  
APPLICANT: Blatner, Frederick R.  
Burland, Nicole T.  
Perna, Nicole T.  
Plunkett, Guy  
Welch, Rod  
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Quarles & Brady  
STREET: 1 South Plinckney Street  
CITY: Madison  
STATE: WI  
COUNTRY: US  
ZIP: 53701-2113  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 Inch, 1.44MB storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/453,702B  
FILING DATE: 03-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-Dec-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296, 95017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 251-5000  
TELEFAX: (608) 251-9166  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61663  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 62:  
US-09-453-702B-62  
Alignment Scores:  
Pred. No.: 85 Length: 61663  
Score: 85.50 Matches: 38  
Percent Similarity: 38.86% Conservative: 30  
Best Local Similarity: 21.71% Mismatches: 77  
Query Match: 8.91% Indels: 30  
Gaps: 6  
US-10-082-502-17 (1-182) x US-09-453-702B-62 (1-61663)

```

OY      5  GLLTFLPLAALALEULEUENGLYLALEULEUGLYTHRALATRALARGTSGERGIN 24
Db      37859  GGCTGGCGTCCGACGATTATACGAGTCTTTGTGTGACCAACTTGGGATCAACAGGGGGCAG 37918
OY      25  ASPLAULHSCYSGLYALACYSARGLALEAULALAPGLULEUGLUTRPSLIULE----- 42
Db      37919  ACAGTGCATTTGATCCCGTAAATAAGTGTTAACGGAGGTGAGCTGTCGATGTTGTGT 37978
OY      43  -----ALAGINVAL-----ASPROLYSLYS-THRIEGLMETCI 54
Db      37979  TTCGGGACCGCTGGAATAATCTCGATGATGGTGGCGCATATACAGACAAATATTATACCTGCAC 38038
OY      54  YSERHEARTGLEASPPOSERTHRHSAAGLYASNTYRVALRVALVALGLYARGLYGL 74
Db      38039  GCGCGTGACAGTTCATTACCGGATCGCCTACGGCGACACAGTGGCTGGCAGTACCCCTGTA 38098
OY      74  RLULALHLSLEUTHRLGULEULEUGLULIIECYASAPARGMETLYSGUTYRGLYGI 94
Db      38099  AGCTGGATAT-----TACGGAT 38116
OY      94  UGLINLEASPPOSERTHRHSAAGLYASNTYRVALRVALVALGLYARGLYGL 114
Db      38117  GACGATGTGATCCGGGGGAGTATGCGGTTTCG--GTGACGGTGAAGAACGACACTGCTGT 38173
OY      114  USERSEGLULEUAPLEUGLNGLYILAEARGYLEASPERASPILESERGLYTHLEULY 134
Db      38174  CTACGGACGTGGCTATTATAGGGGACCGCAAGTACGTGACGCTCAATATGCTTGTTCG 38233
OY      134  SPHEALCYSGUSERILEVALGLU---GLUTYRGILUAPGLULEULIEGLUPHEPHESE 153
Db      38234  C-----CGCATCTCTGTTGTGAGCTTATACGATTACCCGAGACACTGCTGACAGATTTCG 38284
OY      153  RARGLUALAASPANVALLYSAPLYSLEUCYSERLYSARG 167
Db      38285  CGAGATACAGATAATGTGCTGATGACCTTGCACCTATTCTGT 38327

RESULT 5
US-08-257-073-10
: Sequence 10, Application US/08257073
: Patent No. 5766597
: GENERAL INFORMATION:
: APPLICANT: Paoletti, Enzo
: APPLICANT: de Taisne, Charles
: APPLICANT: Tine, John A.
: TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
: NUMBER OF SEQUENCES: 143
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Curtis, Morris & Safford, P.C.
: STREET: 530 Fifth Avenue, 25th Floor
: CITY: New York
: STATE: New York
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/257,073
: FILING DATE: 09-JUN-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/075,783
: FILING DATE: 11-JUN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/852,305
: FILING DATE: 18-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/672,183
: FILING DATE: 20-MAR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Frommer, William S.

```

```

REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2570
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTMS
INFORMATION FOR SEO ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 5181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-257-073-10

Alignment Scores:
Pred. No.: 2.76 Length: 5181
Score: 85.00 Matches: 36
Percent Similarity: 43.31% Conservative: 32
Best Local Similarity: 22.93% Mismatches: 59
Query Match: 8.85% Indels: 30
DB: 1 Gaps: 7

US-10-082-502-17 (1-182) x US-08-257-073-10 (1-5181)

QY 36 AspGluLeuGluTrpGluIleAlaGlnValAspProLysLysThrIleGlnMetCysLys 55
    ||| ||||| ||| :|: ||| :|: :|:
Db 1192 GATCACCCTTGAATTAGATTATTTTAAAGCAAAAAATMAAAAGATGATACACCT 1251

QY 56 PheArgIleAsnProAspGlySerGlnSerValIleGluValProtyrAlaArgSerGlu 75
    :|: ||| ||| :|: :|: :|: :|: :|:
Db 1252 AAATCAACAAGATCCTACGAATCTGTTCAAATPACCAAAAGTCTTATCCAAATGCTATT 1311

QY 76 AlaHis-----LeuThrGluLeuGluGluIleCys-----AspArgMetLys 90
    :|: :|: |||||:|: :|: :|: :|: :|:
Db 1312 GATATATCCTTACACACTGACTGATATTCATTAATTCATTAGCTGCAAGATAAGTAAAT 1371

QY 91 GluTyrGluGluGlnIleAspProSerThrHisArgLysAsnTyrValArgValGlu 110
    |||||:|: :|: ||| ||| :|: :|: :|:
Db 1372 TCAGATAGGTATTTAATGATCTCTGATACTAAAGAAAATTAATGAAAAATTTTACA 1431

QY 111 ArgAsnGluGlu-----SerSerGluLeuAspLeuGlnGly 122
    ||| ||| :|: :|: :|: :|: :|: :|:
Db 1432 GATATTAAGCAAAAGAAAATATTCATTAAATACATTAAAAACAATGATTACAGAA 1491

QY 123 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGlu 142
    :|: ||||:|: :|: :|: :|: :|: :|: :|:
Db 1492 AAAAAAATTAATC-----ACAAAAAGAACAAAATTAATTAATTAATTAATTA 1536

QY 143 GluTyrGlu-----AspGluLeuIleGlu---PhePheSerArgGlu 155
    :|: ||||| :|: :|: :|: :|: :|: :|:
Db 1537 GATATTAATAAGTCAAAAAAGATTAATGAGAAATTAAGTCAAAAAATTAATGAAATGAA 1596

QY 156 AlaAspAsn-----ValLysAspLysLeuCysSerLysArg 167
    :|: |||| :|: ||| |||||:|: ||| |||
Db 1597 TTATATATATTTTGACAAAGATGTCGATGAATAAATATTCAGTCAGAA 1647

RESULT 6
US-08-961-083-159
Sequence 159, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

```

US-10-082-502-17 (1-182) x US-08-961-083-159 (1-1924)

917 CAGGATGATTACATTAAGACGACGAGAGAGCTGAGTTGGATTAAGGAGCTGATGAACCTT 978



OY 126 AspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluTyrGlu 145  
DB 977 CAAATTAAGTTGCTGATTTA-----GAAAAAGAA 1006  
OY 146 AspGluLeuIleGluPhePheSerArgGluAlaAsp-----AsnVal 159  
DB 1007 ATTAGTAACCTTGAAATATATTACTTGAGGGGCTGATCCTGAAAGATGATCTGCTCTT 1066  
OY 160 LysAspLysLeuLysSerLysArgThrAspLeu 170  
DB 1067 CAAATTAATTAAGTTGCTGATTTA-----GAAAAAGAA 1099

## RESULT 8

US-09-071-035-3  
Sequence 3, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1216 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-3

Alignment Scores:  
Pred. No.: 0.931 Length: 1216  
Score: 81.50 Matches: 41  
Percent Similarity: 41.81% Conservative: 33  
Best Local Similarity: 23.16% Mismatches: 58  
Query Match: 8.49% Indels: 45  
DB: 4 Gaps: 10

US-10-082-502-17 (1-182) x US-09-071-035-3 (1-1216)

OY 3 GYTTPGlyTTPLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTyrAlaArgArg 22  
DB 260 GCGTGTATATGCAATGCT----- 277  
OY 23 SerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGlu-----LeuGluTyr 40  
DB 278 GCACAAAGATGAATGTATGTGACTTAACCATATATGATGATGACACAAATCGGCTGG 337

OY 41 GluIleAlaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 338 AAGATGCA-----GACCAATTCGTGAATGTTAGACGCCCAAAATC----- 385  
OY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisIleThrGlu 80  
DB 386 GACGGCAACAA-----TACGGCATTCATTTAATTAATTCG-----ACAGAA 427  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 100  
DB 428 ATGTATCTATAAAGCTGATTTGTTGAAGAATATAGTGTTGAAGTA---CCGAAACA 484  
OY 101 -----HisArgLysAsnTyrValArgVal 109  
DB 485 TTAGACAAATTAAGAACCTCTTAACAAATTAACAAAAATCCAAACAAAGCTGTT 544  
OY 110 GlyArgAsnGlyLysSerGlu-----LeuAspLeuGlnGlyIle 123  
DB 545 GGTGCTGTTTGACTCGTTAAATTAATCTTTACGCAATTGAAATGAAAAACAAAGCGTT 604  
OY 124 ArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGlu 143  
DB 605 GATTTTAATTAAGACTTACATTTTACACAAAGATTC---CAAGAGTCGTGACCTAT 661  
OY 144 TyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 662 TACCGTATGATGATCGAACAGGTTACTTCCGCACAGCTGTTCAAGATAAA 712

## RESULT 9

US-09-071-035-1  
Sequence 1, Application US/09071035  
Patent No. 6448043  
GENERAL INFORMATION:  
APPLICANT: Gil H. Choi  
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
NUMBER OF SEQUENCES: 496  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071.035  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: A. Anders Brookes  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB369P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1347 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-071-035-1

Alignment Scores:  
Pred. No.: 1.08 Length: 1347  
Score: 81.50 Matches: 41

Percent Similarity: 41.81%  
Best Local Similarity: 23.16%  
Query Match: 8.49%  
DB: 4  
Conservative: 33  
Mismatches: 58  
Indels: 45  
Gaps: 10

US-10-082-502-17 (1-182) x US-09-071-035-1 (1-1347)

OY 3 GYTRPGYTRPLeuAlaLeuLeuGluAlaLeuGluYThrAlaTrpAlaArg 22  
DB 368 GCGTGTATGGAAATGCT----- 405  
OY 23 SerGlnAspLeuHisCysGluAlaCysArgAlaLeuValAspGlu-----LeuGluTrp 40  
DB 406 GCACAAAGTGAATGTTGTGGACTTAAACCAATATATGATGATGACAAATCGCGTGG 465  
OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 466 AAAGATGCA-----GACCAATTCGTGAGATATGTAGACGGCCCAAAATC----- 513  
OY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 514 GACGGCAACAA-----TACGGCAATTCATTTAATAATCG-----ACAGAA 555  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGluGluGluIleAspProSerThr 100  
DB 556 ATGTATATCTAATATGCTGATTTGTGAAATAATGCTGTGAAGTA---CCGAAACAA 612  
OY 101 -----HisArgLysAsnTyrValArgValAl 109  
DB 613 TTAGAGCAATTAAAGACGCTTAAACAAATTTACGAAATAATCCAAAGAAAGTGGT 672  
OY 110 G1YArgAsnGluSerSerGlu-----LeuAspLeuGlnGlyIle 123  
DB 673 GGTCTGCTGTTTGGCTGTTAAATACTAATACGCAATTGGAAAGAAACAAAGCGGT 732  
OY 124 ArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGlu 143  
DB 733 GATTATATTAAGACTAGATTAAACAGCAAGATTC---CAAGAACTGCTGCTAT 789  
OY 144 TyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 790 TACCGTATGTCATCGAAGCAGGTACTCCGACACGCTGCTGATCAATAA 840  
RESULT 10  
US-08-750-152A-1  
Sequence 1, Application US/08750152A  
Patent No. 5977331  
GENERAL INFORMATION:  
APPLICANT: ASAKURA, YOKO  
APPLICANT: KIMURA, EICHIRO  
APPLICANT: ABE, CHIZU  
APPLICANT: KAWAHARA, YOSHIO  
APPLICANT: NAKAMATSU, TSUYOSHI  
TITLE OF INVENTION: ALPHA-KETOGLUTARATE DEHYDROGENASE GENE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT, P.C.  
SAREE: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22152  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,152A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: NORMAN F. OBLON  
REGISTRATION NUMBER: 24,618  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4394 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Brevibacterium lactofermentum  
STRAIN: ATCC13869  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 443..4213  
FEATURE:  
NAME/KEY: -35 signal  
LOCATION: 281..287  
FEATURE:  
NAME/KEY: -10 signal  
LOCATION: 307..312  
FEATURE:  
NAME/KEY: RBS  
LOCATION: 421..428  
FEATURE:  
NAME/KEY: terminator  
LOCATION: 4243..4281  
US-08-750-152A-1  
Alignment Scores:  
Pred. No.: 5.93  
Score: 81.50  
Percent Similarity: 43.97%  
Best Local Similarity: 25.53%  
Query Match: 8.49%  
Length: 4394  
Matches: 36  
Conservative: 26  
Mismatches: 58  
Indels: 21  
Gaps: 8  
US-10-082-502-17 (1-182) x US-08-750-152A-1 (1-4394).  
OY 56 PheArgIleAsnProAspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlu 75  
DB 2669 TTCCACGTCGAATGTCATGATGACCCAGAGCGAGTGTGCGGTGGCCAGCTGGCAACCGAG 2728  
OY 76 AlaHis-----LeuThrGluLeuGluGluIleCysAspArgMetLysGlu 91  
DB 2729 TACCGTCGTGCGTTCGGCAAGACGCTTCATGACCTGCTGTCACCGCTCCGCGGC 2788  
OY 92 TyrGluGluGluIleAspProSer---ThrHisArgLysAsnTyrValArgValAlaGly 110  
DB 2789 CACCAAGGAAGTCGATGATCTTCATGACCCAGCAAGATGATAGAGTCATCACCGGC 2848  
OY 111 ArgAsnGluLysSerSerGluLeu-----AspLeuGlnGlyIleArgIleAspSerAsp 128  
DB 2849 CCGGAGACCGCTTCGTCTCAGTACACCGAAGACGCTCGGA-----CGTGAAGAC 2899  
OY 129 IleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyrGluAspGlu 148  
DB 2900 CTCTCAACGAA-----GATCAGAAAGAGTCGTCGCGGACTTCACAGCAAGATG 2950  
OY 149 IleGluPhePheSerArgGluAlaAspAsnValLysAspLysLeuCysSerLysArgThr 168  
DB 2951 GAATCTGTTCACAGGAAGTCAAGGAGCGCGCAAG---AAGCAGGCTGAGAGCAGACAC 3007  
OY 169 AspLeu-----CysAspHisAlaLeu-----HisIleSerHisAspGlu 181  
DB 3008 GGCATCACCGGCTCCCAAGACCTTCACACGCGCTTGAGACCAACATCTCCGTAAGAG 3067



QY 149 -----11egluPheSerArgGluAlaAspAsnValLysAsp 161  
Db 4389643 GCCCATGACATCGCGGTGTGTGTCACAGTCACGCGGGAGTGTG-----GAC 4389596  
QY 162 LysLeuGlySerLysArgThrAspLeuGlyAspHisAla 174  
Db 4389595 GCGTGTGCTGCGAGATCATCCCGGTGTGTGTGCGACGCC 4389557  
RESULT 13  
US-08-837-029-1  
Sequence 1, Application US/08837029  
Patent No. 5945303  
GENERAL INFORMATION:  
APPLICANT: Mel et al.  
TITLE OF INVENTION: Human Hematopoietic - Specific Protein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/837,029  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US96/04930  
FILING DATE: 11-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Brooks, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF268  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 833 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 42..608  
FEATURE:  
NAME/KEY: s19\_peptide  
LOCATION: 42..107  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 108..608  
US-08-837-029-1  
Alignment Scores:  
Pred. No.: 0.623 Length: 833  
Score: 81.00 Matches: 46  
Percent Similarity: 37.81% Conservative: 30  
Best Local Similarity: 22.89% Mismatches: 81  
Query Match: 8.44% Indels: 44  
Gaps: 10  
US-10-082-502-17 (1-182) x US-08-837-029-1 (1-833)  
QY 7 LeuAlaLeuLeuGlyAlaLeu-----LeuGly-----16  
Db 63 CTGCTGCTGCTGCTGAGAGCTGAGCATCCAGAGGGGCGCTGGGGAGACGGCGCCACTC 122

QY 17 ThrAlaThrAlaArgSerGlnAsp-----25  
Db 123 ACAGCCACACCTCCACACTGATGAGGATGTACTACGCCACATGCCCGCTCAC 182  
QY 26 LeuHisGlyGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGluIleAlaGlnVal 45  
Db 183 CTGCGCTGTGATGCTGACAGAGCTGTGCTTACAGATGTGGCAAAATCTGGCAAAAGCA 242  
QY 46 AspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer 65  
Db 243 GAGACCAAACTTCATACCTCA-----AACTTGGGGGGCGCGGAA 284  
QY 66 ValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluIle 85  
Db 285 CTGACCGAGTTGGTCTAC-----ACGATGTCTTGACCGGAC 323  
QY 86 CysAspArg---MetLysGluTyrGly---GluGlnIleAspProSerThrHisArgLys 103  
Db 324 TGCTCCCGCACTGCGACGACTACGAGTTCTGAGAGTTGAGAGTTGAC-----365  
QY 104 AsnTyrValArgValValGlyArgAsn---GlyGluSerSerGluLeuAspLeuGlnGly 122  
Db 366 CAACTGAACCTCTCACAGGCCAGACTTACGAGGGGCCAGACCACTACGCTG 425  
QY 123 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGlu 142  
Db 426 ATGTCTACAGGGGGCCCTGCGCTTACCAGAGCTTCCAGACATGTTGCATCTTGGGG 485  
QY 143 GluTyr---GluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAsp 161  
Db 486 GAGTTTGAGAGAACGATCTATGAAAGCCACCAACAAAGCCGAGGGCTTGGAGGCA 545  
QY 162 LysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisIleSerHisAspGlu 181  
Db 546 TTGCTATGTGGGGAGACCCAGGGGCTGCTCAGAGAGGTGTACGACCAAGAGAGAG 605  
QY 182 Leu 182  
Db 606 CTC 608  
RESULT 14  
US-08-993-380-3  
Sequence 3, Application US/08993380B  
Patent No. 6077676  
GENERAL INFORMATION:  
APPLICANT: Shl, Otimel  
TITLE OF INVENTION: SINGLE CHAIN POLYPEPTIDES COMPRISING TROPONIN I AND  
FILE REFERENCE: 1112-1-053  
CURRENT APPLICATION NUMBER: US/08/993,380B  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 3  
LENGTH: 1173  
TYPE: DNA  
ORGANISM: Homo sapiens (modified)  
US-08-993-380-3  
Alignment Scores:  
Pred. No.: 1.57 Length: 1173  
Score: 79.50 Matches: 47  
Percent Similarity: 35.41% Conservative: 27  
Best Local Similarity: 22.49% Mismatches: 58  
Query Match: 8.28% Indels: 77  
Gaps: 8  
US-10-082-502-17 (1-182) x US-08-993-380-3 (1-1173)  
QY 10 LeuLeuGlyAlaLeuLeuGlyThrAlaThrAlaArgArgSerGlnAspLeuHisGlyGly 29  
Db 10 LeuLeuGlyAlaLeuLeuGlyThrAlaThrAlaArgArgSerGlnAspLeuHisGlyGly 29

```

Db 460 ATGATGACAGGGCTGCTGGG---GCCGGGGCTAAGAGACTCCCTGAGCTG----- 507
Oy 30 A1ACysArgAlaLeuValAspLeuGluTyrGluIleAlaGluValAspProIysIys 49
Db 508 -----CGGGCCACCCTCAAGCAGGTGAAGAGAGACACCAG-----AAGGAA 552
Oy 50 ThrIleGlnMetGlySerPheArgIleAsnProAsp----- 61
Db 553 AACCGGAGGTGGAGACTGGCGCAAGAACATGATGACAGTACAGTGAATGAGAGGCCGC 612
Oy 61 ----- 61
Db 613 AAGAAAAGTTTGAGAGACACATGAGTGGTGGTGTCTGCTGGGGGGGCTTCTGTGGC 672
Oy 62 GlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeu 81
Db 673 GGTGTGTTCATCATGATGATGATCATCTACAGAGCTCCGCTAGACACAGTACAGAAAG 732
Oy 82 -----LeuGluGluIleCys 86
Db 733 CAGAAAATGACTTCAAGGACACCTTCGACATCTTCGTGCGGCGGTGAGAGTGGCTGC 792
Oy 87 AsparGlnMetGlyGluTyrGlyGluIleAsnProSerThrHisArgIysAsnTyrVal 106
Db 793 ATCAGCACCAAGAGAGCTGGGCAAG-----GTGATG 822
Oy 107 ArgValValGlyArgAsnGlyCysLeuSerGluLeuAspLeuGluIleArgIleAsp 126
Db 833 AGGATGCTGGGGCGAGAACCCCACTGAGGAGCTGCAGAGATGATCATGATGAGTGGAC 882
Oy 127 SerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleVal----- 141
Db 883 GAGGACGCGCAGGCGCAGCTGGACTTTGAT---GAGTTCCTGTGATGATGCTGGGTC 939
Oy 142 -----GluGluTyrGluAspLeuIleGluPhe 152
Db 940 ATGAGAGACAGACGAAAGGAATCTGAGAGAGAGCTGTCTGACCTCTCCGACATGTTT 999
Oy 153 SerArgGluAlaAspAsnValIysAsp 161
Db 1000 GACAAAATGCTGATGGCTTACATCGAC 1026

RESULT 15
US-08-938-105-2
Sequence 2, Application US/08938105
Patent No. 6353151
GENERAL INFORMATION:
APPLICANT: Leinwand, Leslie A.
APPLICANT: Vikstrom, Karen L.
TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver
STATE: CO
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938.105
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crook, Wamell M.
REGISTRATION NUMBER: 31,071
REFERENCE/DOCKET NUMBER: 3595-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700

```

```

; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5661 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5661
; US-08-938-105-2

Alignment Scores:
Pred. No.: 20.2 Length: 5661
Score: 78.50 Matches: 37
Percent Similarity: 43.02% Conservative: 40
Best Local Similarity: 20.67% Mismatches: 71
Query Match: 8.18% Indels: 31
DB: Gaps: 7

US-10-082-502-17 (1-182) x US-08-938-105-2 (1-5661)
Oy 17 ThrAlaItrAlaArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAsp 36
Db 2656 ACGGCCAAGACAGCGCAAGCTGGAGAC---GAGTGTCTCAGAGCTCAAGAAAGATTCGAT 2712
Oy 37 GlnLeuGluTyrGluIleAlaGlnValAspProIysIysThrIleGlnMetGlySerPhe 56
Db 2713 GACCTGAGAGCTACACCTGGCCCAAGGTGGAGAAAGAAACACAGCAAGAAAGAGT 2772
Oy 57 ArgIleAsnProAspGlySerGlnSerValValGluVal-----ProTyrAlaArgSer 74
Db 2773 AAAAAGCTGACAGAGAGATGGCCGGCGCTGCAGAGATCATTCGCAAGCTGCACCAAGAG 2832
Oy 75 GlnAlaHisLeuThrGluLeuGluGluIleCysAspArgMetIysGluTyrGlyGlu 94
Db 2833 AAGAAAGCTTTCAGAGAGGCCCAACGACCAAGCCCTAGATGACCTTCAGGCTGAGAGAGC 2892
Oy 95 GlnIleAspProSerThrHisArgIysAsnTyrValArgValValGlyArgAsnGlyGlu 114
Db 2893 AAGGTCAACACACTGACC-----AAGTCTAAGTCAAGCTG-----GAG 2931
Oy 115 SerSerGluLeuAspLeuGlnGly-----IleArgIleAspSer 127
Db 2932 CAGCAGGTGGATGATCTGGAGGATCCCTGAGCAGCAGCAAGAGATGGCGCATGACCTG 2991
Oy 128 Asp-----IleSerGlyThrLeuLysPheAlaCysGluSerIleValGlu 142
Db 2992 GAGCGAGCAAGCGAGAGCTGGAGGCTGACCTGAAAGCTGACCCAGAGAGATCATGGAC 3051
Oy 143 GlnTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValIysAspIys 162
Db 3052 CTGGAAGACAGCAAGCTT-----CAGCTGAGGAGAAAG 3084
Oy 163 LeuCysSerIysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspLeu 181
Db 3085 CTCAGAGAGAAAGAGTTTGATCATCATCAGCAGCAAGACAGTAAATAGAGAGCAGCAG 3141

Search completed: March 31, 2003, 04:04:16
Job time : 1662.66 secs

```

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 07:02:46 ; Search time 55.0233 Seconds

(without alignments)  
2814.757 Million cell updates/sec

Title: US-10-082-502-17

Perfect score: 960  
Sequence: 1 MKGGMALALGALGTAMA.....LCSKRDLDHALHSHDEL 182

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Zgapop 6.0 , Zgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame\_plus\_p2n\_model -DEV=slp  
-O=/cgn2\_1/USPTO.spool/US10082502/runat\_24032003\_135103\_6870/app.query.fasta\_1.1308  
-DB=Published.Applications\_NA -QFMT=fastcap -SUFFIX=trpb -MIMATCH=0.1  
-LOOPT=0 -LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62  
-TRANS=human40.cst -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR\_MAX=100  
-TRN\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NOR=ext -HEATSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10082502\_6CGN\_1.1.145\_6runat\_24032003\_135103\_6870  
-NCPU=6 -ICPU=3 -NO\_XLPRX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications\_NA:\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PTI_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PTCIS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	960	100.0	814	10	US-09-864-864-309
2	960	100.0	1210	9	US-10-028-072-173
3	960	100.0	1210	9	US-10-121-049-173
4	960	100.0	1210	9	US-10-123-904-173

5	960	100.0	1210	9	US-10-140-470-173	Sequence 173, App
6	960	100.0	1210	9	US-10-175-746-173	Sequence 173, App
7	960	100.0	1210	9	US-10-176-918-173	Sequence 173, App
8	960	100.0	1210	9	US-10-176-921-173	Sequence 173, App
9	960	100.0	1210	9	US-10-137-865-173	Sequence 173, App
10	960	100.0	1210	9	US-10-140-474-173	Sequence 173, App
11	960	100.0	1210	9	US-10-142-831-173	Sequence 173, App
12	960	100.0	1210	9	US-10-143-114-173	Sequence 173, App
13	960	100.0	1210	9	US-10-140-002-173	Sequence 173, App
14	960	100.0	1210	9	US-10-142-819-173	Sequence 173, App
15	960	100.0	1210	9	US-10-123-262-173	Sequence 173, App
16	960	100.0	1210	9	US-10-142-823-173	Sequence 173, App
17	960	100.0	1210	9	US-10-121-050-173	Sequence 173, App
18	960	100.0	1210	9	US-10-141-755-173	Sequence 173, App
19	914	95.2	657	10	US-09-822-846-398	Sequence 398, App
20	266	27.7	564	10	US-09-864-864-228	Sequence 228, App
21	131.5	13.7	1457	10	US-09-846-573B-10	Sequence 10, App1
22	131.5	13.7	1470	9	US-10-036-041-58	Sequence 58, App1
23	131.5	13.7	1470	9	US-10-035-855-58	Sequence 479, App
24	131.5	13.7	1470	9	US-10-174-590-479	Sequence 479, App
25	131.5	13.7	1470	9	US-10-176-758-479	Sequence 479, App
26	131.5	13.7	1470	9	US-10-175-737-479	Sequence 479, App
27	131.5	13.7	1470	9	US-10-173-706-479	Sequence 479, App
28	131.5	13.7	1470	9	US-10-175-738-479	Sequence 479, App
29	131.5	13.7	1470	9	US-10-175-752-479	Sequence 479, App
30	131.5	13.7	1470	9	US-10-176-482-479	Sequence 479, App
31	131.5	13.7	1470	9	US-10-176-757-479	Sequence 479, App
32	131.5	13.7	1470	9	US-10-176-757-479	Sequence 479, App
33	131.5	13.7	1470	9	US-10-176-913-479	Sequence 479, App
34	131.5	13.7	1470	9	US-10-180-552-479	Sequence 479, App
35	131.5	13.7	1470	9	US-10-180-557-479	Sequence 479, App
36	131.5	13.7	1470	9	US-09-931-836-58	Sequence 58, App1
37	131.5	13.7	1470	9	US-10-173-700-479	Sequence 479, App
38	131.5	13.7	1470	9	US-10-174-572-479	Sequence 479, App
39	131.5	13.7	1470	9	US-10-174-579-479	Sequence 479, App
40	131.5	13.7	1470	9	US-10-174-588-479	Sequence 479, App
41	131.5	13.7	1470	9	US-10-174-588-479	Sequence 479, App
42	131.5	13.7	1470	9	US-10-175-739-479	Sequence 479, App
43	131.5	13.7	1470	9	US-10-175-740-479	Sequence 479, App
44	131.5	13.7	1470	9	US-10-175-743-479	Sequence 479, App
45	131.5	13.7	1470	9	US-10-176-488-479	Sequence 479, App

#### ALIGNMENTS

RESULT 1  
US-09-864-864-309  
Sequence 309, Application US/09864864  
Patent No. US20020102679A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Secrist, Heather  
APPLICANT: Lodes, Michael J.  
APPLICANT: Algate, Paul A.  
APPLICANT: Fling, Steve P.  
APPLICANT: Mannion, Jane  
APPLICANT: Benson, Darin R.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.523  
CURRENT APPLICATION NUMBER: US/09/864,864  
CURRENT FILING DATE: 2001-05-23  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: Corlita Invention Disclosure Database  
SEQ ID NO 309  
LENGTH: 814  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-864-309





;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/066278  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069334  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/072320  
;; PRIOR FILING DATE: 1998-01-23  
;; PRIOR APPLICATION NUMBER: 60/073612  
;; PRIOR FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28

;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 19/98-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090663  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

## Alignment Scores:

Pred. No.: 2,54e-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00 Conservative: 0  
Best Local Similarity: 100.00 Mismatches: 0  
Query Match: 100.00 Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-028-072-173 (1-1210)

QY 1 MettysGLYTrpGLYTrpLeuAlaLeuLeuGLYAlaLeuLeuGLYThrAlaTrpAla 20  
DB 543 ATGAAGGCTGGGCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602  
QY 21 ArgatSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspLeuLeuGluTrp 40  
DB 603 CGAGAGAGCCAGGATGCTCCAGTGTGAGCATGAGGCTCTGTGTGATGAATGAG 662  
QY 41 GluLeuAlaGluValAspProLysIsthrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTTGCCAGGTGAGCCCAAGCAACATTGATGATGATGATGATGATGATGATG 722  
QY 61 AspGlySerGlnSerValAlaGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 723 GATGCGACGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
DB 783 CTGCTGAGAGAGATATGTGACCGGATGAAGAGATGAGGAAACATGATGATGATGATG 120  
QY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 140  
DB 843 CATGCCAAGAACTACGTACGTGTAGTGGCCCGAATGAGATCCAGTAAGTGAACCTGACCT 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATCGAATGCACTCAATATTTAGCGGACCCCTCAAGTTTGCCTGTGAGACATT 962

OY 141 ValGIuGIuTYrGIuAspGIuLeuILEGIuPheSerArgIuAlaAspAsnValIys 160  
DB 963 GTGGAGATGCGAGGAGTCACTGATTCCTTTCCCGAGAGGCTGACAAATGTATAA 1022  
OY 161 AspIysIeuCysSerIysArgThrAspIeuCysAspHisAlaLeuHisILeserHisasp 180  
DB 1023 GACAAACTTTTCAGTAAACCAACAGATCTTTGTGACATGCCCTGCACATATCGCATAT 1082  
OY 181 GluLeu 182  
DB 1083 GAGCTA 1088

## RESULT 3

US-10-121-049-173  
; Sequence 173, Application US/10121049  
; Publication No. US20030022239A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C17  
; CURRENT APPLICATION NUMBER: US/10/121,049  
; CURRENT FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 550  
; Prior Application removed - See File Wrapper or Palm  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-121-049-173

## Alignment Scores:

Pred. No.: 2.54e-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-121-049-173 (1-1210)

OY 1 MetIysGIuTYrGIuTYrPLeuAlaLeuIeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGCTGCTGGGGGCGCTGCGGAACGGCTGGGCT 602  
OY 21 ArgArgSerGIuAspIeuHisCysGIuAlaCysArgAlaLeuValAspGIuLeuGIuTrp 40  
DB 603 CGGAGGAGCCAGGATCTCCACTGTGAGCATGCAAGGCTGTGGATGAGTGAATGG 662  
OY 41 GIuIleAlaGIuValAspArgTrpIysThrILeGIuMetGIuSerPheArgIleAsnPro 60  
DB 663 GAAATTCGCCAGGTGGACCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722  
OY 61 AspIysSerGIuSerValValGIuValProTYrAlaArgSerGIuAlaHisLeuThrGIu 80  
DB 723 GATGGACGCGGATGAGTGGAGTGGCTTATGCCCGCTCAGAGGCCCACTGACAGAG 782

OY 81 LeuLeuGIuGIuIleCysAspArgMetIysGIuTYrGIuGIuGIuInIleAspProSerThr 100  
DB 783 CTGCTGGAGGAGATATGTGACCGGATGAGGAGTATGGGAAACAGATGATCTTCCACC 842  
OY 101 HisArgIysAsnTYrValArgValValGIuYrArgAsnGIuIuSerSerGIuLeuAspLeu 120  
DB 843 CATCGCAAGAACTACGTACGTGTAGTGGCCGGAATGGAGAAATCCAGTGAACCTGACCTA 902  
OY 121 GIuGIuIleArgIleAspSerAspIleSerGIuThrLeuIysPheAlaCysGIuSerIle 140  
DB 903 CAAGCATCCCAATGCATGCATGATTTAGCGGACCCCTCAAGTTTGCCTGTGAGAGCAT 962  
OY 141 ValGIuGIuTYrGIuAspGIuLeuILEGIuPheSerArgIuAlaAspAsnValIys 160  
DB 963 GTGGAGATGCGAGGAGTCACTGATTCCTTTCCCGAGAGGCTGACAAATGTATAA 1022  
OY 161 AspIysIeuCysSerIysArgThrAspIeuCysAspHisAlaLeuHisILeserHisasp 180  
DB 1023 GACAAACTTTTCAGTAAACCAACAGATCTTTGTGACATGCCCTGCACATATCGCATAT 1082  
OY 181 GluLeu 182  
DB 1083 GAGCTA 1088

## RESULT 4

US-10-123-904-173  
; Sequence 173, Application US/10123904  
; Publication No. US20030022328A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C54  
; CURRENT APPLICATION NUMBER: US/10/123,904  
; CURRENT FILING DATE: 2002-04-16  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-123-904-173

## Alignment Scores:

Pred. No.: 2.54e-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-123-904-173 (1-1210)

OY 1 MetIysGIuTYrGIuTYrPLeuAlaLeuIeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGCTGCTGGGGGCGCTGCGGAACGGCTGGGCT 602

```
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 603 CGAGGAGCCAGGAGGATCTCCACTGTGAGCATGAGGCGCTTGTTGGATGAACTAGAAATGG 662
QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60
Db 663 GAATATGCCAGGTGGATCCCAAGAAAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80
Db 723 GATGGCAGCCAGTCACTGGTGAGTGCCCTTATGCCCGCTCAGAGGCCACCTCAGACAGAG 782
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 783 CTGCTGAGAGATATGTGATCCGATGAAAGAGATATGGGAAAGATGATCTTCCCTCCACC 842
QY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 120
Db 843 CATCGCAAGACATCACTAGCTGTAGTGGCGCGGAAATGGAAATCCACTGAACTGGACCTA 902
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGGCATCCGAAATCAGTCACTAGATATTAGCGCGACCCCTCAAGTTGGCTGAGAGCAT 962
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 963 GTGGAGAAATACGAGATGTAACATCTTAATCTTTCCGAGAGCGCTGCAATGTGTTAA 1022
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
Db 1023 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATGCGATGAT 1082
QY 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 5
US-10-140-470-173
; Sequence 173, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-470-173

Alignment Scores: 2.54e-113 Length: 1210
Pred. No.: 1
```

```
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-140-470-173 (1-1210)
QY 1 MetLysGlyTyrProGlyTyrPheAlaLeuLeuGlyAlaLeuGluGlyThrAlaTrpAla 20
Db 543 ATGAAAGCGTGGGTTGGCTGGCCCTCTTCTGTGGGGCCCTGCTGGAAACCGCTGGGCT 602
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 603 CGAGGAGCCAGGATCTCCACTGTGAGCATGAGGCGCTTGTTGGATGAACTAGAAATGG 662
QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60
Db 663 GAATATGCCAGGTGGATCCCAAGAAAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80
Db 723 GATGGCAGCCAGTCACTAGCTGTAGTGGCGCGGAAATGGAAATCCACTGAACTGGACCTA 902
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 783 CTGCTGAGAGATATGTGATCCGATGAAAGAGATATGGGAAAGATGATCTTCCCTCCACC 842
QY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 120
Db 843 CATCGCAAGACATCACTAGCTGTAGTGGCGCGGAAATGGAAATCCACTGAACTGGACCTA 902
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGGCATCCGAAATCAGTCACTAGATATTAGCGCGACCCCTCAAGTTGGCTGAGAGCAT 962
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 963 GTGGAGAAATACGAGATGTAACATCTTAATCTTTCCGAGAGCGCTGCAATGTGTTAA 1022
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
Db 1023 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATGCGATGAT 1082
QY 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 6
US-10-175-746-173
; Sequence 173, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
```

		CURRENT APPLICATION NUMBER: US/10/175,746	
		CURRENT FILING DATE: 2002-06-19	
		Prior Application removed - See File Wrapper or Palm	
		NUMBER OF SEQ ID NOS: 550	
		SEQ ID NO 173	
		LENGTH: 1210	
		TYPE: DNA	
		ORGANISM: Homo Sapien	
		US-10-175-746-173	
Alignment Scores:			
Pred. No.:		2,54e-113	Length: 1210
Score:		960.00	Matches: 182
Percent Similarity:		100.00%	Conservative: 0
Best Local Similarity:		100.00%	Mismatches: 0
Query Match:		100.00%	Indels: 0
DB:		9	Gaps: 0
US-10-082-502-17 (1-182) x US-10-175-746-173 (1-1210)			
QY	1	MettySGlyTTPGlyTTPLeuAlaLeuLeuGlyAlaLeuLeuGlyThraLarPaLa	20
Db	543	ATGAAAGCTGGGGTTGGCTGGCCCTCTTGGGGCCCTGCTGGAACTCCCTGGGCT	602
QY	21	ArgArSergInaAspLeuHnHSrGSyGlyAlaCyArGAlaLeuValaAspLluLeuGluTrp	40
Db	603	CGAGAGACCAGATCTCCACTGTGAGACATCGAGGGCTCTGGTGATTAACCTAGAAATGG	662
QY	41	GlulLeaGluValaAspProLysTrhLleGlnMetGlySerPheArGllLeaAnPro	60
Db	663	GAATTCGCCCGAGTGGACCCCAAGAACACATTCACATGGATCTTCCGATCAATCA	722
QY	61	AspGlySergInSerValValGluValProTrYAlaArgSergLualHnISleuThrglu	80
Db	723	GATGGCAGCCAGTCAGTGGTGGAGGTGCCCTTATGCCCTCAGAGGCCACCTCACAGAG	782
QY	81	LeuLeuGluGluTrGlyLeuCyAspArGMetLysGluTrGlyGluGlnLeaPProSerThr	100
Db	783	CTGTGGAGAGATATGTGACCGGATGMAAGGATATAGGGGAACAAGTTGATCTTCCACC	842
QY	101	HIsArGlyAsnTrYValaArgValValGlyArgAsnGlyGluSerSergLluLeuAspLeu	120
Db	843	CATCGCAAGAACAACAGTACGTGATAGTGGGCCGGAATGGAGATCCAGTGAACCTGACCTA	902
QY	121	GlnGlyLleArGllLeaAspSerApIleSerglyThrLeuLysPheAlaCysGluSerIle	140
Db	903	CMAAGCATCCGAATCGACTCGATGATATTAGGGGACCCCTCAAAATTGGCGTGAGAGCAATT	962
QY	141	ValGluGluTrGlyGluAspGluLeuIleGluPheSerArgGluLualAspAsnValLys	160
Db	963	GTGAGAGAAATACGAGATGAGTCACTTGAATCTTTTCCCGAGAGGCTGACAAATGTTAA	1022
QY	161	AspLysLeuCySerLysArgTrhAspLeuCyAspPheAlaLeuHnISleSerHAsp	180
Db	1023	GACAAACTTTCAGATTAAGCAACAGATCTTTTGAGACAAATGCCCTGCACATATCGCATGAT	1082
QY	181	GluLeu 182	
Db	1083	GAGCTA 1088	
RESULT 7			
US-10-176-918-173			
Sequence 173, Application US/10176918			
Publication No. US20030027275A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin P.			
APPLICANT: Beresini, Maureen			
APPLICANT: DeForge, Laura			
APPLICANT: Desnoyers, Luc			
APPLICANT: Filvaroff, Ellen			
APPLICANT: Gao, Wei-Qiang			
APPLICANT: Gerritsen, Mary E.			
APPLICANT: Goddard, Audrey			

APPLICANT:	Godowski, Paul J.		
APPLICANT:	Guiney, Austin L.		
APPLICANT:	Sherwood, Steven		
APPLICANT:	Smith, Victoria A.		
APPLICANT:	Stewart, Timothy A.		
APPLICANT:	Tunas, Daniel		
APPLICANT:	Watanabe, Colin K		
APPLICANT:	Wood, William		
APPLICANT:	Zhang, Zemin		
TITLE OF INVENTION:	SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC		
TITLE OF INVENTION:	ACIDS ENCODING THE SAME		
FILE REFERENCE:	P33081C82		
CURRENT APPLICATION NUMBER:	US/10/176, 918		
CURRENT FILING DATE:	2002-06-20		
Prior Application removed - See File Wrapper or Palm			
NUMBER OF SEQ ID NOS:	550		
SEQ ID NO 173			
LENGTH:	1210		
TYPE:	DNA		
ORGANISM:	Homo Sapien		
US-10-176-918-173			
Alignment Scores:			
Pred. No.:	2,54e-113		
Score:	960.00		
Percent Similarity:	100.008		
Best Local Similarity:	100.008		
Query Match:	100.00%		
DB:	9		
US-10-082-502-17 (1-182) x US-10-176-918-173 (1-1210)			
QY	1	MellysglytrpjltrpLleuAlaleuLeuGlYAlaleuLeuGlYThralatrpAla	20
Db	543	ATGAAGAGCTGGGGTGGCTGGCCCTCTCTGGGGGCCCTGCTGGAAACCGCTGGCT	602
QY	21	ArgatgserGlnaspLeuHIScysGlyAlAcysArgAlaleuValaspGluLeuGluTrp	40
Db	603	CGGAGGAGCCAGATCTCCACTGCTGGAGCATGcAGGGCTCTGCTGATGAACTAGAAATG	662
QY	41	GluIleuAlagluValaspProlylsVsthrILeGImetGlyserPheArgIleaspPro	60
Db	663	GAATTGGCCAGGTGGACCCCAAGACACATTCAGATGGGATCTTCCGGATCAATCCA	722
QY	61	AspGlyserGlnserValalGluValProtyrAlaArgserGluAlaHisLeuThrGlu	80
Db	723	GATGGCAGCCAGTcAGTGGTGGAGGCTCCTTATGCCCGCTCAAGGCCCACTCAAGAG	782
QY	81	LeuLeuGluGluIleCysAspArgMetLysGluTrpGlyGluGluIleAspProserThr	100
Db	783	CTGTGGGAGGAATGTGACCGGATGMAAGAGATAGGGGACACAGATTGCTTCACCC	842
QY	101	HisArgLysAsnTrpValaArgValalValArgAsnGlyGluSerGluLeuAspLeu	120
Db	843	CATGGCAAGAACTcGTAAGTGTAGTGGCCGGAATGGAGAAATCCAGTGAAGTGGACCTA	902
QY	121	GlnGlyIleArgIleaspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	140
Db	903	CAAGCAGATCCGAATCGACTCAGATATATAGGGGACACCTCAAGTTTCCGTGTAGACATT	962
QY	141	ValGluGluTrpGluaspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys	160
Db	963	GTGGAGGAATACGAGATGTGAACATCATGATTTCTTTCCGAGAGGCTGACATATGTAA	1022
QY	161	AspLysLeuLysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp	180
Db	1023	GACAAACTTTGCAAGTAAAGCCAAAGATCTTTGTGACCATGACCTGACATATTCGATCAT	1082
Db	1083	GAGCTA 1088	

```
US-10-176-921-173
; Sequence 173, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Tumas, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C28
; CURRENT APPLICATION NUMBER: US/10/176,921
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-082-502-17 (1-182) x US-10-176-921-173 (1-1210)
QY 1 MetlysglytrpglytrpleualaleuleuleuglyalaleuleuglythrAlaTrpAla 20
Db 543 ATGAAGGCGGGGTGGCTGGCCCTCTCTGGGGGCGCTGGGAACCGCTGGGCT 602
QY 21 ArgArSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGGCTGTGGTGAAGTGAAGTGAAGTGA 662
QY 41 GluIleAlaGlnValAspProlyslsYThrIleGlnMetClySerPheArgIleAspPro 60
Db 663 GAAATTGCCAGGTGGAGCCCAAGAGACCATTCAGATGGGATCTTCCGATCAATCCA 722
QY 61 AspGlySerGlnSerValValAlaGluValProTyAlaArgSerGluAlaHisLeuThrGlu 80
Db 723 GATGGAGCCAGTCAGTGGGTGGAGTGCCTTATGCCCCGCTCAAGGCGCCACTCACAGAG 782
QY 81 LeuLeuGluGluIleCysAspArgMetIleGluTyTrGlyGluGlnIleAspProSerThr 100
Db 783 CTGCTGGAGAGATATGTGACCCGATGAGAGATATGGGACACAGATGTGCTTCACCC 842
QY 101 HisArgLyAsnTyValAlaArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 120
Db 843 CACCGGAGAACTACTACTGTAGTGGCCGGAATGGAATCAAGTGAACAGTGAGACTCA 902
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAAGGATCCGAATCACTCAGATATTAAGGGCACCCTCAAGTTTCACTGTGAGACCAT 962
QY 141 ValGluGluIleCysAspArgMetIleGluIlePheSerArgIleAlaAspAsnValLys 160
```

```
Db 963 GTGAGAGATACGAGATGACATCTTCAATTTCTTCCGAGAGGCTGACAAATGTTAA 1022
QY 161 AsplysIleCysSerIleArgTyThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
Db 1023 GACAAACTTTCAGTAAAGCAAGATCTTGTGACCATGCGCTGGACATATTCGATAT 1082
QY 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 9
US-10-137-865-173
; Sequence 173, Application US/10137865
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-865-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-082-502-17 (1-182) x US-10-137-865-173 (1-1210)
QY 1 MetlysglytrpglytrpleualaleuleuleuglyalaleuleuglythrAlaTrpAla 20
Db 543 ATGAAGGCGGGGTGGCTGGCCCTCTCTGGGGGCGCTGGGAACCGCTGGGCT 602
QY 21 ArgArSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGGCTGTGGTGAAGTGAAGTGAAGTGA 662
QY 41 GluIleAlaGlnValAspProlyslsYThrIleGlnMetClySerPheArgIleAspPro 60
Db 663 GAAATTGCCAGGTGGAGCCCAAGAGACCATTCAGATGGGATCTTCCGATCAATCCA 722
QY 61 AspGlySerGlnSerValValAlaGluValProTyAlaArgSerGluAlaHisLeuThrGlu 80
Db 723 GATGGAGCCAGTCAGTGGGTGGAGTGCCTTATGCCCCGCTCAAGGCGCCACTCACAGAG 782
QY 81 LeuLeuGluGluIleCysAspArgMetIleGluIlePheSerArgIleAlaAspAsnValLys 100
```

```
DB 783 CTGCTGAGACATATATGTACCGGATGAGAGATATGGGAGACAGATTGATCTTCACC 842
QY 101 H1A9GLYASNTYrValArqValValGlyArqAsnGlyGlySerSerGluLeuAspLeu 120
DB 843 CATCGCAAGACTACTAGTGTAGTGGCCGGAATGGAATCCAGTCACTGAGACCTA 902
QY 121 GINGLYTleArGTLleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
DB 903 CAAGGATCCGATCGATCACTACATATATTACCGGACCCCTCAAGTTGCGTGTGAGACATT 962
QY 141 ValGluGluTyrGluAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnValLys 160
DB 963 GTGGAGGAATACGAGCATATCACTGTAATTTCTTTCCCGAGAGGCTGACAAATGTTAA 1022
QY 161 AspLysLeuCysSerLysArGThrAspLeuCyAspPheAlaLeuHisIleSerHisAsp 180
DB 1023 GACAAACTTTGCAAGTAAAGCAAGACAGATCTTTGTGACCAATGCCCTGACAAATCGCATGAT 1082
QY 181 GuluLeu 182
DB 1083 GAGCTA 1088

RESULT 10
US-10-140-474-173
; Sequence 173, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-082-502-17 (1-182) x US-10-140-474-173 (1-1210)
QY 1 MetLysGlyTyrGlyTyrPheuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrrpAla 20
DB 543 ATGAAAGCGTGGGCTGGCTGGCTGCTTGTGGGGGCCCTGCGAAGCCGCTGGGCT 602
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrp 40
```

```
DB 603 CGAGAGAGCAGGATATCTCCAGTGTGAGCATCGAGGCTCTGTGGATGAACATAGATGG 662
QY 41 GuluLeuAlaGluValAspProLysTyrIleGluMetGlySerPheArgIleAsnPro 60
DB 663 GAAATTCGCCAGGTGGAGCCCAAGAAAGACCATTCACATGGATCTTTCCGATCAATCCA 722
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
DB 723 GATGGACAGCAGTCAAGTGTGGAGGCTTATGCCCCCTCAGAGGCCACCTCACAGAG 782
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
DB 783 CTGCTGAGAGCATATGTACCGGATGAAAGAGATATGGGAGACAGATTGATCTTCACC 842
QY 101 H1A9GLYASNTYrValArqValValGlyArqAsnGlyGlySerSerGluLeuAspLeu 120
DB 843 CATCGCAAGACTACTAGTGTAGTGGCCGGAATGGAATCCAGTCACTGAGACCTA 902
QY 121 GINGLYTleArGTLleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
DB 903 CAAGGATCCGATCGATCACTACATATATTACCGGACCCCTCAAGTTGCGTGTGAGACATT 962
QY 141 ValGluGluTyrGluAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnValLys 160
DB 963 GTGGAGGAATACGAGCATATCACTGTAATTTCTTTCCCGAGAGGCTGACAAATGTTAA 1022
QY 161 AspLysLeuCysSerLysArGThrAspLeuCyAspPheAlaLeuHisIleSerHisAsp 180
DB 1023 GACAAACTTTGCAAGTAAAGCAAGACAGATCTTTGTGACCAATGCCCTGACAAATCGCATGAT 1082
QY 181 GuluLeu 182
DB 1083 GAGCTA 1088

RESULT 11
US-10-142-431-173
; Sequence 173, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173

Alignment Scores:
Pred. No.: 2,54e-113 Length: 1210
Score: 960.00 Matches: 182
```

```

/ CURRENT FILING DATE: 2002-05-09
/ Prior Application removed - See Palm or File Wrapper
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 173
/ LENGTH: 1210
/ TYPE: DNA
/ ORGANISM: Homo Sapien
US-10-143-114-173

Alignment Scores:
Pred. No.:      2,54e+113          Length:      1210
Score:          960.00             Matches:      182
Percent Similarity: 100.00%       Conservative: 0
Best Local Similarity: 100.00%     Mismatches:  0
Query Match:    100.00%           Indels:      0
DB:              Gaps:            0

US-10-082-502-17 (1-182) x US-10-143-114-173 (1-1210)

QY      1 MetLysGIrTPgIYTrrPLeuAlaleuLeuencLYlAleuLeuGlyThrAlaTrPaLa 20
Db      543 ATGAAGCGCTGGGTGGTGGCCCTGCCTTGggGGGCCCTGCGGGACCGCCTGGGCT 6020
QY      21 ArgArSeriGlnAsPLeuHnAScYsgLYAlAcSaRgaLAleuValAsPgLuLeuGLutrp 40
Db      603 CGGAGGACCcAGATrTCtCAcTGrGGAGcANtGCAGGGcCTcGTGGATGAACTrGAAntGG 6620
QY      41 GLullealagInvalAspproLYsLystrThrIlegImetCLyserphearGileasPro 60
Db      663 GAATTGGCCcAGtgGgAACCCCMAGAMACCATTCAGATGGGATCTTCCGGATCAATCCA 7220
QY      61 AspCLySerGInserValAlgluValProtyrAlaArgSerGluAlhISleuThrGlu 80
Db      723 GATGGcAGCcAGCTAgGTgrGTGAGAGTGCCTTAtrccCGCTCAGAGGCCcACTCACAGAG 7820
QY      81 LeuLeucGLugluILecYsaSPArGaMeLlysGlutryGluGlnIlleasProSerThr 100
Db      783 CTCTGTGAGGAGATATGTGACCGGATGAAGAGATGGGGAACAGATGATCTTCCTCAC 8420
QY      101 HisArgLYsaAnTYrrVALargValVAlglyARgAnsclYguIserserGluLeuAspleu 120
Db      843 CATGGcAGNANCtRAGrACGTGrTAGTGGGCCGGAATGGAAATCCAGTGMACTGGACCTA 9020
QY      121 GlnGLylleaRgIlleasSerAsPIleserGlyThrLeuLYSPhealAcYSgIusertlle 140
Db      903 CAAGcATCCGAATCGACtCAcCATGATATAGGGCGCACCTCAAGTTTGCGTGTGAGACATT 9620
QY      141 ValGLugluLyrrGluAspgLUleuIlleGluPheserArGclunIAAspaAnaVallys 160
Db      963 GTGGAGGAAATACGAGGTATGAACCATGATTCTTTCCGAGAGGCTGACANTGTATAA 1020
QY      161 AsPLyLSerCYsserLYSArGTHrAsPLeuCyaSPShAlaleuHstIleserHisasp 180
Db      1023 GACAACATTTCAGcGTAAAGcAACAGATCTTGTGACCATGCCTCGACATATCGCATGAT 1080
QY      181 Gluleu 182
Db      1083 GAGCTA 1088

RESULT 13
US-10-140--002-173
/ Sequence 173, Application US/10140002
/ Publication No. US20030037623A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowsky4, Paul T
```



APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P330R1C59  
CURRENT APPLICATION NUMBER: US/10/140,002  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-002-173

Alignment Scores:  
Pred. No.: 2,546-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-140-002-173 (1-1210)

OY 1 MetLysGlyTyrGlyTyrPleuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrrPa 20  
DB 543 ATGAAAGCGTGGGTGGCTGGCCCTGCTTCTGGGGGCCCTGGGAACCCCTGGGCT 602  
OY 21 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrP 40  
DB 603 CGGAGAGCCAGGATCTCCACTGTGGAGCATGCAAGGCTCTGTGGATGATAGATGG 662  
OY 41 GlnLeuAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCGGATCAATCCA 722  
OY 61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
DB 723 GATGGCAGCAGCATGAGTGGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAG 782  
OY 81 LeuLeuGlnGluIleCysAspArgMetLysGluTyrGlyGlnGlnIleAspProSerThr 100  
DB 783 CTGCTGGAGAGATATGTACCCGATGAAAGAGTATGGGGAACAGATTGATCTTCACCC 842  
OY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLusSerSerGluLeuAspLeu 120  
DB 843 CATCCAGAACTACGATAGTGTAGTGGGCCGGAATGGAAATCCAGTGAATGGAGCTTA 902  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATTCGCAATGACACAGATATTTAGCCGACCTCAACTTTCCTGTAGAGCAT 962  
OY 141 ValGlnGluTyrGlnAspGluLeuIleGluPhePheSerArgGlnAlaAspAsnValLys 160  
DB 963 GTGGAGAAATACAGAGATGAACTCATTTGAAATCTTTCCCGAGAGGCTGACATGTAAA 1022  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
DB 1023 GACAAACTTTGAGATAGAGAAACAGATCTTTGTGACCATGCCCTGCACATATGCAATG 1082  
OY 181 GluLeu 182  
DB 1083 GAGCTA 1088  
RESULT 14  
US-10-142-419-173

Sequence 173, Application US/10142419  
GENERAL INFORMATION:  
US2003004945A1  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P330R1C244  
CURRENT APPLICATION NUMBER: US/10/142,419  
CURRENT FILING DATE: 2002-05-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-142-419-173

Alignment Scores:  
Pred. No.: 2,546-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-17 (1-182) x US-10-142-419-173 (1-1210)

OY 1 MetLysGlyTyrGlyTyrPleuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrrPa 20  
DB 543 ATGAAAGCGTGGGTGGCTGGCCCTGCTTCTGGGGGCCCTGGGAACCCCTGGGCT 602  
OY 21 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrP 40  
DB 603 CGGAGAGCCAGGATCTCCACTGTGGAGCATGCAAGGCTCTGTGGATGATAGATGG 662  
OY 41 GlnLeuAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCGGATCAATCCA 722  
OY 61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
DB 723 GATGGCAGCAGCATGAGTGGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAG 782  
OY 81 LeuLeuGlnGluIleCysAspArgMetLysGluTyrGlyGlnGlnIleAspProSerThr 100  
DB 783 CTGCTGGAGAGATATGTACCCGATGAAAGAGTATGGGGAACAGATTGATCTTCACCC 842  
OY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLusSerSerGluLeuAspLeu 120  
DB 843 CATCCAGAACTACGATAGTGTAGTGGGCCGGAATGGAAATCCAGTGAATGGAGCTTA 902  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATTCGCAATGACACAGATATTTAGCCGACCTCAACTTTCCTGTAGAGCAT 962  
OY 141 ValGlnGluTyrGlnAspGluLeuIleGluPhePheSerArgGlnAlaAspAsnValLys 160  
DB 1083 GAGCTA 1088  
RESULT 14  
US-10-142-419-173

DB 963 GTGAGCAATACGAGATGATGATGATTCTTTCCCGAGAGGCTGACAAATGTTAAA 1022  
QY 161 AsPLysLeuCySerLySArGThrAsPLeuCyAsPHisAlaLeuHisIleSerHisAsp 180  
DB 1023 GACAAACTTGGAGTAAGGACAGATCTTTGTGACCATGCCCTGCACATATGCAATGAT 1082  
QY 181 GluLeu 182  
DB 1083 GAGCTA 1088  
RESULT 15  
US-10-123-262-173  
; Sequence 173, Application US/10123262  
; Publication No. US20030049816A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C38  
; CURRENT APPLICATION NUMBER: US/10/123,262  
; CURRENT FILING DATE: 2002-04-15  
; Prior Application removed - See file wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-123-262-173  
Alignment Scores:  
Pred. No.: 2,54e-113 Length: 1210  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-10-082-502-17 (1-182) x US-10-123-262-173 (1-1210)  
QY 1 MetLysGlyTTPGLYTRPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThraLarTPala 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCCCTGCTCTGGGGCCCTGCTGGGAACCGCTGGGCT 602  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40  
DB 603 CGAGAGAGCCAGATCTCCACTGTGAGCATGCGAGGCTCTGTGATGAACAGATG 662  
QY 41 GluIleAlaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAAATTTGCCAGGTGACACCCCAAGAACACCATTCACATGGATCTTCCGATCAATCCA 722  
QY 61 AspgLySerGlnSerValValGluValProTyraLarGserGluAlaHisLeuThrGlu 80  
DB 723 GATGGACGACAGTCACTGTGTGAGTGGCTTATGCCGCTCAGAGGCCACCTCACAGAG 782  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 100

DB 783 CTGCTGAGAGAGATATANGTACCCGATGAAGAGACTATGGGGAACAGATTGATCTTCACC 842  
QY 101 HisArgLysAsnTyraValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 843 CATGGCAAGACTAGCTAGCTGTAGTGGCCGGAATGGAAATCCAGTGAACCTGGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATCCGAATGCACTCAGATATTAGCGCCACCCCTCAAGTTGCGTGTGAGACATT 962  
QY 141 ValGluGluTyrgluAspGluLeuIleGluPhePheSerArgLysAlaAspAsnValLys 160  
DB 963 GTGAGCAATACGAGATGAACTCATTTCTTTCCCGAGAGGCTGACAAATGTTAAA 1022  
QY 161 AsPLysLeuCySerLySArGThrAsPLeuCyAsPHisAlaLeuHisIleSerHisAsp 180  
DB 1023 GACAAACTTGGAGTAAGGACAGATCTTTGTGACCATGCCCTGCACATATGCAATGAT 1082  
QY 181 GluLeu 182  
DB 1083 GAGCTA 1088

Search completed: March 31, 2003, 11:59:04  
Job time : 58.0233 secs



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:43:09 ; Search time 1166.6 Seconds

(without alignments)  
2526.643 Million cell updates/sec

Title: US-10-082-502-17

Perfect score: 960  
Sequence: 1 MKGCGWLLALLGALLGTA.....LCSKRTDLCDAHLHSHDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame+pn.model -DEV=xlp  
-O/cgn2.1/USPTO.spool/US10082502/runat\_24032003.135100.6599/app.query.fasta\_1.1308  
-DB=EST -OEXT=fastap -SUFFIX=st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdt -LIST=45  
-DOCALIGN=200 -THR\_SCORE=Pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTEXT=Pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502@cgn2.1.5425@runat\_24032003.135100.6599 -NCPD=6 -ICPD=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-NARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl1:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_luv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	960	100.0	588	14	BM832828	BM832828 K-EST0107
2	960	100.0	591	14	BM832888	BM832888 K-EST0107
3	960	100.0	611	13	BT172899	BT172899 1d98f09.y
4	960	100.0	697	12	BE344334	BE344334 602014708
5	960	100.0	719	12	BE733900	BE733900 601568370
6	960	100.0	742	12	BC824549	BC824549 602728454
7	960	100.0	745	12	BE791763	BE791763 601581586
8	960	100.0	1117	13	BM552907	BM552907 AGENCOURT
9	954	99.4	880	14	BQ421827	BQ421827 AGENCOURT
10	952	99.2	783	13	BT550436	BT550436 603192905
11	947	98.6	868	14	BQ421069	BQ421069 AGENCOURT
12	945	98.4	791	12	BT550443	BT550443 602688633
13	939	97.8	773	13	BT550162	BT550162 603193045
14	938	97.7	754	11	AK013568	AK013568 Mus muscu
15	938	97.7	772	11	AK007914	AK007914 Mus muscu
16	938	97.7	808	13	BT408525	BT408525 602964935
17	938	97.7	843	13	BT412539	BT412539 602990995
18	938	97.7	853	11	AK013014	AK013014 Mus muscu
19	938	97.7	1255	11	AK019927	AK019927 Mus muscu
20	934	97.3	724	14	BQ210416	BQ210416 UT-R-DY1-
21	934	97.3	818	12	BE870315	BE870315 601447464
22	934	97.3	875	13	BT763950	BT763950 603049829
23	933	97.2	887	13	BT762759	BT762759 603048045
24	927	96.6	673	14	BQ668530	BQ668530 AGENCOURT
25	927	96.6	720	13	BT909759	BT909759 603070841
26	926	96.5	802	12	BF608664	BF608664 MYL 00160
27	925	96.4	891	13	BT408381	BT408381 602964013
28	920	95.8	783	13	BT763784	BT763784 603049729
29	920	95.8	914	12	BE799495	BE799495 601589289
30	911	94.9	765	12	BF312281	BF312281 601898804
31	906	94.4	866	12	BF965088	BF965088 602268856
32	901	93.9	665	10	AW919569	AW919569 EST350873
33	900	93.8	866	13	BT454180	BT454180 603170602
34	899	93.6	992	13	BT456021	BT456021 603170466
35	898	93.5	751	12	BE742698	BE742698 601574753
36	892	92.9	709	13	BT411532	BT411532 602965126
37	887	92.4	797	13	BT762927	BT762927 603047987
38	883	92.0	885	14	BQ920767	BQ920767 AGENCOURT
39	883	92.0	963	14	BQ929020	BQ929020 AGENCOURT
40	880	91.7	908	14	BQ877139	BQ877139 AGENCOURT
41	880	91.4	1019	12	BG024176	BG024176 602303060
42	877	91.4	776	13	BT522523	BT522523 603175203
43	874	91.0	699	9	AL548119	AL548119 AT548119
44	869	90.5	746	13	BT551339	BT551339 603193144
45	867	90.3	930	14	BQ437660	BQ437660 AGENCOURT

ALIGNMENTS

RESULT 1  
LOCUS BM832828  
DEFINITION K-EST0107375 S55NU484s1 Homo sapiens cDNA clone S55NU484s1-11-B03  
ACCESSION BM832828  
VERSION BM832828.1 GI:19189237  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 588)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

TITLE  
JOURNAL  
COMMENT

FEATURES

Kim, Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsungemail.kr@kribb.re.kr  
Plate: 11 row: B column: 03  
High quality sequence stop: 588.  
Location/Qualifiers

BASE COUNT

149 a 140 c 176 g 123 t

Alignment Scores:

Pred. NO.: 1.75e-111 Length: 588  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-082-502-17 (1-182) x BM832828 (1-588)

QY 1 MettysglttprgltyttrleuAlaleuLeuGClYThralatrpAla 20  
Db 2 ATGAAGGCTGGGGTGGCTGGCTCTGCTGGGGCCCTGCGAACCCTGGGCT 61  
QY 21 ArgatgserglnaapleuHnscysgltYalacysargalaleuValaspclueuLutrp 40  
Db 62 CGGAGACCCAGGATCTCCACTGTGGACAGTCAAGGGCTGTGGATGAATGAATGG 121  
QY 41 GlutlealaglnValaspProLysThrIleGlnMetGlySerPheArgIleasnPro 60  
Db 122 GAATATGCCAGGTGGACCCCAAGAACCAATTCATGATGGGATCTTTCCGGATCAATCA 181

QY 61 AspglyserGlnSerValaIgluValProTyrAlaArgSerGlnAlaHisLeuthrglu 80  
Db 182 GATGCAGCCAGTCACTGATGGATGGCTTATGCCCGCTCAGAGGCCACCTCAGAG 241  
QY 81 LeuLeuGluGluIleCysaspParMetysgltYtYrGlyGluGlnIleaspProserThr 100  
Db 242 CTGCTGAGAGATATGTATCCAGATGAGAGAGTATGGGGAACAGATTCATCTTCACAC 301  
QY 101 HisArgLysAsnTyrValaArgValaIglYArgAsnGlyGluSerSerGluLeuaspLeu 120  
Db 302 CATCCAGAACTACGATACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 361  
QY 121 GlnGlyIleArgIleaspSeraspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 362 CAAGCATCTCCAGATGAGTCACTAGATATGAGCGCACCTCAAGTTCGGTGAAGACAT 421  
QY 141 ValGluGluTyrGluaspGluLeuIleGluPhePheSerArgGluAlaAspAsnValys 160  
Db 422 GTGAGGAATACGAGATGAACTCATGAAATCTTTCCCGAGAGGCTGCACAAATGTAA 481  
QY 161 AspLysLeuCyserLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisasp 180  
Db 482 GACAAACTTTCAGATGAGGACGACAGATCTTTGTGACATGCCCTGCACATTCGATGAT 541  
QY 181 GluLeu 182  
Db 542 GAGCTA 547  
RESULT 2  
BM832888  
LOCUS  
DEFINITION  
K-EST0107445 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-11-H01  
5', mRNA sequence.  
ACCESSION  
BM832888  
VERSION  
BM832888.1 GI:19189297  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 591)  
Kim, Y.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
Oh, K.U., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
Kim, Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsungemail.kr@kribb.re.kr  
Plate: 11 row: H column: 01  
High quality sequence stop: 591.  
Location/Qualifiers

FEATURES

source  
1. 591  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="S5SNU484s1-11-H01"  
/clone\_1lb="S5SNU484s1"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A) + RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dT-selected  
mRNA by priming with dT-tailed vector. The cDNA vector  
was adjusted to have about 60nt. The cDNA vector

was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(dT)<sub>14</sub> as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-961. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by the transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method.

BASE COUNT 150 a 142 c 176 g 123 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1,76e-111 Length: 591  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-17 (1-182) x BM632888 (1-591)

QY 1 MetLysGlyTTPGlyTTPLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTTPala 20  
DB 2 ATGAAAGGCTGGGGTGGGCTGGCCCTGCTCTGGGGCCCTGGGAAACGGCTGGGCT 61  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40  
DB 62 CGAGAGAGCCAGATCTCCACTGTGAGCATGAGGCTTGGTGGATGAGTAAAGATGG 121  
QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 122 GAATATGCCAGGTGGACCCCAAGAACCATTCAGATGGAGATCTTCGGAGCAATCCA 181  
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
DB 182 GATGGAGCCAGTCACTGTGGAGTGCCCTTATGCCCGCCAGAGGCCACCTCACAGAG 241  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr 100  
DB 242 CTCTGTGAGAGATATGTGACCGGATGAGAGATATGGGAGACAGATTGATCCTTCACC 301  
QY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 302 CATCCGAAAGAACTACGTACGTAGTGGGCCGGAATGGAATCCAGTGAACCTGAGACC 361  
QY 121 GlnGlyIleArgGlnLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 362 CAAGGATCCGAATGCACTCAGATATATAGCGGACCCCTCAAGTTTCCGTGTGAGACAT 421  
QY 141 ValGluGluTyrGlnAspGluLeuIleGluPhePheSerArgGlnAlaAspAsnValLys 160  
DB 422 GTGGAGGAATACGAGATGATGAACTTCTTTCCGAGAGGCTGACAAATGTTTAA 481  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180  
DB 482 GACAAACTTTCAGTAAAGGAGACAGATCTTTGACACATCCCTGCACATATGCATGAT 541  
QY 181 GluLeu 182  
DB 542 GAGCTA 547

RESULT 3  
B1712899

LOCUS B1712899 611 bp mRNA linear EST 11-MAR-2002  
DEFINITION id98f09.y1 HR85 15let Homo sapiens cDNA clone IMAGE:5085832 5'

similar to TR:09Y2B0 Q9Y2B0 TYPE II MEMBRANE PROTEIN. ; mRNA  
Sequence: B1712899

ACCESSION B1712899.1 GI:15688594  
VERSION EST.

KEYWORDS human.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 611)

Mellon,D., Brown,J., Kenty,G., Pernutt,A., Lee,C., Kaestner,R.,  
Lemishka,I., Scaecce,M., Brestelli,J., Gradwohl,G., Clifton,S.,  
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,  
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas  
,M., Gibbons,M., McManus,R., Cole,R., Tsagariswilli,R., Williams,T.,  
, Jackson,T., and Bowers,Y.

Endocrine Pancreas Consortium  
Unpublished (2000)

CONTACT: Douglas Mellon, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138

Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmellon@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For Information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco  
High quality sequence stop: 463.

Location/Qualifiers  
1. 611

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5085832"  
/clone\_lib="HR85 15let"  
/tissue\_type="Purified pancreatic 15let"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size -1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

BASE COUNT 157 a 143 c 184 g 127 t  
ORIGIN

## FEATURES

Alignment Scores:  
Pred. No.: 1,85e-111 Length: 611  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-17 (1-182) x B1712899 (1-611)

QY 1 MetLysGlyTTPGlyTTPLeuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTTPala 20  
DB 30 ATGAAAGGCTGGGGTGGGCTGGCCCTGCTCTGGGGCCCTGGGAAACGGCTGGGCT 89  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40  
DB 90 CGAGAGAGCCAGATCTCCACTGTGAGCATGAGGCTTGGTGGATGAGTAAAGATGG 149  
QY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60

|||||  
DB 150 GAAATTGCCAGGTGAGCCCAAGAGACCAATTCAGATGGATCTTCCGAGTCAATCCA 209  
OY 61 ASpglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 210 GATGGCAGCCAGTCACTAGTGTGAGAGTGCCTTATGCCCGCTCAGAGGCCACTCCACAGAG 269  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 270 CTGCTGGAGGAGATATGTGACCGGATGAAGAGATGAGGGAACAGATGATCTTCCACCC 329  
OY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 330 CATCCCAAGAACTACGTACGTGTAGTGGACCGGAATGGAATCCAGTAATGAGACTGAACCTA 389  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 390 CAAGCATCCGAAATGAGTCAAGATATAGCGGCACCCCTCAAGTTGGCTGTAGAGCAATT 449  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 450 GTGGAGGATACGAGCACTCAATCTTCTTCCGAGAGCGCTGACAAATGTATAA 509  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheIleAlaLeuHisIleSerHisAsp 180  
DB 510 GACAAACTTTGAGTACGAGAACAGATCTTGTGACACATGCCCTGCACATATGCATGAT 569  
OY 181 GluLeu 182  
DB 570 GAGCTA 575  
RESULT 4  
BF344334 697 bp mRNA linear EST 22-NOV-2000  
LOCUS BE344334  
DEFINITION 602014708F1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4150512  
5', mRNA sequence.  
ACCESSION BE344334  
KEYWORDS BE344334.1 GI:11291554  
SOURCE EST.  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 697)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
plate: LLM9413 row: P column: 01  
High quality sequence stop: 692.  
Location/Qualifiers  
1. .697  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4150512"  
/clone\_id="NCI\_CGAP\_Brn64"  
/tissue\_type="globlastoma with EGFR amplification"  
/lab\_host="DH10B (TI phage-resistant)"  
/note="Organ: brain; Vector: PCMV-SPORT; Site\_1: NCI;  
Site\_2: Salt; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.57 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 173 a 167 c 220 g 137 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,25e-111 Length: 697  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
US-10-082-502-17 (1-182) x BF344334 (1-697)  
OY 1 MetLysGlyTyrPglTyrPLeuAlaLeuLeuGluGlyAlaLeuGluGlyThrAlaTrrPLe 20  
DB 92 ATGAAAGCTGGGGTGGCTGGCTGGCTTGTGGGGGCTGTGGGAAACCGCGGGCT 151  
OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyrP 40  
DB 152 CGAGAGACCGAGAACTCCACTGTGTGAGACATCCAGCGCTGTGTGAGTGAATGAGTGAATGG 211  
OY 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60  
DB 212 GAAATTGCCAGGTGAGCCCAAGAGACCAATTCAGATGGATGATCTTCCGATCAATCCA 271  
OY 61 ASpglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 272 GATGGCAGCCAGTCACTAGTGTGAGAGTGCCTTATGCCCTCAGAGCCCTCAGACAG 331  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 332 CTGCTGGAGGAGATATGTGACCGGATGAAGAGATAGGGGAACAGATGATCTTCCACCC 391  
OY 101 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 392 CATCCCAAGAACTACGTACGTGTAGTGGCGGGAATGGAATCCAGTAATGAGACTGAGACTA 451  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 452 CAAGCATCCGAAATGAGTCAAGATATAGCGGCACCCCTCAAGTTGGCTGTAGAGCAATT 511  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 512 GTGGAGGATACGAGCACTCAATCTTCTTCCGAGAGCGCTGACAAATGTATAA 571  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheIleAlaLeuHisIleSerHisAsp 180  
DB 572 GACAAACTTTGAGTACGAGAACAGATCTTGTGACACATGCCCTGCACATATGCATGAT 631  
OY 181 GluLeu 182  
DB 632 GAGCTA 637  
RESULT 5  
BE733900 719 bp mRNA linear EST 15-SEP-2000  
LOCUS BE733900  
DEFINITION 6015568370F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3843209 5',  
mRNA sequence.  
ACCESSION BE733900  
KEYWORDS BE733900.1 GI:10147892  
SOURCE EST.  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 719)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be



found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
plate: L1CM537 row: k column: 18  
high quality sequence stop: 710.  
Location/Qualifiers

## FEATURES

source

1. 719  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3843209"  
/clone\_lib="NIH\_MGC\_21"  
/issue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: placenta; Vector: pORF7; Site.1: XhoI;  
Site.2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(g). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 190 a 164 c 214 g 151 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2,35e-111 Length: 719  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-082-502-17 (1-182) x BE733900 (1-719)

OY 1 MetlySGlyTTPpGlyTTPleuAlaleuLeuenglyAlaleuenglyThraATpPa 20  
Db 48 ATGAAAGCGTGGGTGGCGCCCTCTCTGGGGCCCTGCGGAAACCGCTGGGCT 107  
OY 21 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValaspGluLeuGluTTP 40  
Db 108 CGAGAGAGCCAGATCTCCCTCTGTGAGCATGCAAGGCTGTGTGATGATGAATGAG 167  
OY 41 GluileAlaGlnValaspProLysLysThrileGlnmetGlySerPheArgileasPro 60  
Db 168 GAATATGCCAGGTGAGCCCAAGAAAGACCATTCAGATGGATCTTCCGATCAATCCA 227  
OY 61 AspGlySerGlnSerValValaspGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 228 GATGGAGCCAGTCAGTGGTGGAGGTGCTTATGCCGCTCAGAGGCCACCTCACAGAG 287  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGluIleAspProSerThr 100  
Db 288 CTGCTGAGAGATATGTGACCCGATGAAAGAGTATGGGAAACAGATTGATCTTCACAC 347  
OY 101 HisArgLysAsnTyrValaArgValValaspGlyArgasnGlyLysSerGluLeuAspLeu 120  
Db 348 CATCGCAAGAACTACGACGTAGTGGGCCGGAATGGAATCCAGTGAACAGGACCTA 407  
OY 121 GlnGlyIleArgGlyLeaspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 408 CAAGGATCCGAAATCAGATCAATTAAGGCGACCCCTCAAGTTGGCTGTGAGAGCAT 467  
OY 141 ValGluGluTyrGluaspGluLeuileGluPhePheSerArgGluAlaAspAsnValLys 160  
Db 468 GTGAGAGATACAGAGATGACATCAATTTCTTTCCCGAAGGCTGCAATGTATAA 527  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 180  
Db 528 GACAAACTTGGCAGTAAGCAAGACAGATCTTGTGACCATGCCCTGCACATATGCATGAT 587  
OY 181 GluLeu 182  
Db 588 GAGCTA 593  
RESULT 6

BG824549

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 742)

NIH-MGC

http://mgc.ncl.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1CM1737 row: k column: 13

high quality sequence stop: 740.

Location/Qualifiers

1. 742

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4868100"

/clone\_lib="NIH\_MGC\_15"

/issue\_type="adenocarcinoma cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pORF7; Site.1: XhoI; Site.2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(g). Size-selected &gt;500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 180 a 184 c 233 g 144 t

ORIGIN

## Alignment Scores:

Pred. No.: 2,47e-111 Length: 742  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-082-502-17 (1-182) x BG824549 (1-742)

OY 1 MetlySGlyTTPpGlyTTPleuAlaleuLeuenglyAlaleuenglyThraATpPa 20  
Db 141 ATGAAAGCGTGGGTGGCGCCCTCTCTGGGGCCCTGCGGAAACCGCTGGGCT 200  
OY 21 ArgArgSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValaspGluLeuGluTTP 40  
Db 201 CGAGAGAGCCAGATCTCCCTGTGAGCATGCAAGGCTGTGTGATGATGAATGAG 260  
OY 41 GluileAlaGlnValaspProLysLysThrileGlnmetGlySerPheArgileasPro 60  
Db 261 GAATATGCCAGGTGAGCCCAAGAAAGACCATTCAGATGGATCTTCCGATCAATCCA 320  
OY 61 AspGlySerGlnSerValValaspGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 321 GATGGAGCCAGTCAGTGGTGGAGGTGCTTATGCCGCTCAGAGGCCACCTCACAGAG 380  
OY 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGluIleAspProSerThr 100

Db 381 CTGCTGAGAGATATGTGACCGGATGAAGAGATGCGGAACAGATTGATCCTTCACAC 440  
Oy 101 HIsAgLySaNTyrValArgValAlaIylArgAsnGlyGluSerSerGluLeuAspLeu 120  
Db 441 CATCCCAAGAACTAGCTAGTGTAGTGGCCGGAATGGAATCCAGAACTGACCTTA 500  
Oy 121 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 501 CAGGCAATCCGAAATGACTGAGATATAGCGGACCCCTCAAGTTGGCTGCTAGAGCAT 560  
Oy 141 ValGluGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 561 GTGAGGAATACGAGAGATGAACTCATTTCTTTCCGAGAGCGTACAAATGTTAA 620  
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 180  
Db 621 GACAAACTTTCAGTACGAAAGCAAGATCTTGTGACCATGCGCTGCACATATGCGATGAT 680  
Oy 181 GluLeu 182  
Db 681 GAGCTA 686

RESULT 7  
BE791763 745 bp mRNA linear EST 20-SEP-2000  
LOCUS BE791763  
DEFINITION 601581586F1 NIH\_MGC\_7 Homo sapiens CDNA clone IMAGE:3936141 5',  
mRNA sequence.  
ACCESSION BE791763  
VERSION BE791763.1 GI:10212961  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 745)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LCM779 row: k column: 22  
High quality sequence stop: 739.  
Location/Qualifiers  
1..745  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3936141"  
/clone\_1id="NIH\_MGC\_7"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOTB7; site\_1: XhoI; site\_2:  
EcoRI; CDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adapter: GGGACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 182 a 180 c 237 g 146 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,48e-111 Length: 745  
Score: 960.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-10-082-502-17 (1-182) x BE791763 (1-745)

Oy 1 MetLysGlyTyrGlyTyrLeuAlaLeuLeuGluGlyAlaLeuGlyThrAlaTyrAla 20  
Db 138 ATGAAAGCTGGGGTGGCTGGCCCTGCTCTGGGGGCCCTGCTGGAAACCGCTGGGCT 197  
Oy 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlyTyr 40  
Db 198 CGAGAGACCGAAGATCTCCACTGTGAGACATGACAGGCTGTGGTGAATGAATGAATGG 257  
Oy 41 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgTLeuAsnPro 60  
Db 258 GAATATGCCAGCTGGAGCCCAAGAAAGACATTCAGATGGGATCTTCCGGATCAATCCA 317  
Oy 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 318 GATGGCAGCCAGTCAGTGTGGAGGTGCTTATGCCCCCTCAGAGGCCCTCAGACAG 377  
Oy 81 LeuLeuGluGluIleCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
Db 378 CTGCTGAGAGATATGTGACCGGATGAAAGAGTATGGGAAACAGATTGATCTTCACAC 437  
Oy 101 HIsAgLySaNTyrValArgValAlaIylArgAsnGlyGluSerSerGluLeuAspLeu 120  
Db 438 CATCCCAAGAACTAGCTAGTGTAGTGGCCGGAATGGAATCCAGAACTGACCTTA 497  
Oy 121 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 498 CAGGCAATCCGAAATGACTGAGATATAGCGGACCCCTCAAGTTGGCTGCTAGAGCAT 557  
Oy 141 ValGluGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 558 GTGAGGAATACGAGAGATGAACTCATTTCTTTCCGAGAGCGTACAAATGTTAA 617  
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 180  
Db 618 GACAAACTTTCAGTACGAAAGCAAGATCTTGTGACCATGCGCTGCACATATGCGATGAT 677  
Oy 181 GluLeu 182  
Db 678 GAGCTA 683

RESULT 8  
BM552907 1117 bp mRNA linear EST 20-FEB-2002  
LOCUS BM552907  
DEFINITION AGNCOURT\_6572552 NIH\_MGC\_41 Homo sapiens CDNA clone IMAGE:5466942  
5', mRNA sequence.  
ACCESSION BM552907  
VERSION BM552907.1 GI:18791172  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1117)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: LCM1968 row: c column: 07  
High quality sequence stop: 709.  
Location/Qualifiers

FEATURES

```

source
1. .1117
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone_image="546842"
   /clone_lib="NIH_MGC_41"
   /tissue_type="amelanotic melanoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI. cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH_MGC library."
BASE COUNT      299 a      258 c      320 g      239 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:          4,53e-111           Length:          1117
Score:             960.00              Matches:         182
Percent Similarity: 100.00%            Conservative:    0
Best Local Similarity: 100.00%          Mismatches:     0
Query Match:       100.00%             Indels:         0
DB:                Gaps:               0

US-10-082-502-17 (1-182) x BM552907 (1-1117)
OY      1 MetLeuGIYTrpGLYTrpLeuAlaLeuLeuLuuGLYAlaLeuLeuGLYThrAlaTrpAla 20
        |||||||
Db      148 ATCGAAGCCTGGGGCTGGCTGGCTGGCTGGGGGCCCTGGTGGAACCGCCGTGGGCT 207
OY      21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlnTrp 40
        |||||
Db      208 CGGAGGAGCCAGAGATCTCCACTGTGAGCATCAAGGCTGTGTGATGAATCAATGG 267
OY      41 GluIIleAlaGluValAspProIysIysThrlIegImetGlySerPheArgILEasPro 60
        |||||||
Db      268 GAATTGGCCAGAGTGGACCCTCAAGAACCAATTCAGATGGGATCTTCCGATCAATCCA 327
OY      61 AspGlySerGlnSerValValaGluValaProTYrAlaArgSerGluAlaHisLeuThnGlu 80
        |||||
Db      328 GATGGCAGCACTCACTGTGTGGAGGTGGCTTAATGCCCTCAGAGGCCACTCACAGAG 387
OY      81 LeuLeuGluGluILIEcysApArgMetIysGluTrpGlyGluGlnIleAspProSerThr 100
        |||||
Db      388 CTGGCTGGAGGAGAATATGTGACCGGATGAAGAGATATGGAGAAAGATTTGATCCAC 447
OY      101 HisArgLyAsnTYrValaArgValaGluTyArgsnGlyGluSerSerGluLeuAspLeu 120
        |||||
Db      448 CATGCCAAGAACTACGTACCTGTACTGTGGGCGGAATGAGAACTGCAATCTGGACCTA 507
OY      121 GInglyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
        |||||
Db      508 CAAGCATTCCCAATTCGACTCAGATATATAGCGCAACCTCAAGTTTGGCTGTGAGACGATT 567
OY      141 ValIGluGluTYrGluAspGluLeuIleGluPheheserArgGluAlaAspAsnValLys 160
        |||||
Db      568 GTGGAGGAATACGAGAGATGAACATACTGAATCTTTTCCCAGAGGCTGACAATGTGTAAA 627
OY      161 AspLysLeuCySserLYsArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp 180
        |||||
Db      628 GACAAACTTTGCAGTAGCAGAACGAAAGATCTTTGTACCATGCCCTGCACATATGCATGAT 687
OY      181 Glueu 182
        ||||||
Db      688 GAGCTA 693

```

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
BQ421827	1	GI:21117142	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	NIH-MGC	http://mhc.nci.nih.gov/	1 (bases 1 to 880)	Unpublished (1999)
					Contact: Robert Strausberg, Ph. D.		Email: c99bbs-ref@mail.nih.gov		Tissue Procurement: ATCC/DCMP/DRP
					cDNA Library Preparation: Life Technologies, Inc.		DNA Sequencing by: Agencourt Bioscience Corporation		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://limage.llnl.gov">http://limage.llnl.gov</a>
					Plate: L14M13303		row: a	column: 01	High quality sequence stop: 580.
FEATURES					Location/Qualifiers				
source					1..880				
					/organism="Homo sapiens"				
					/db_xref="taxon:9606"				
					/clone="IMAGE:6050568"				
					/clone_id="NIH.MGC.72"				
					/issue_type="melanotic melanoma"				
					/lab_host="DH10B (phage-resistant)"				
					/note="Organ: skin; Vector: PCMV-SPOrt6; Site:1; Noti: Site:2; SalI; Cloned unidirectionally. Primer: Oligo dT Average insert size 2 kb. Library constructed by Life Technologies."				
BASE COUNT	238 a	208 c	256 g	176 t	2 others				
ORIGIN									
Alignment Scores:									
Score: No.:	1.84e-110	Length:	880						
Percent Similarity:	954.00	Matches:	181						
Best Local Similarity:	99.45%	Conservative:	0						
Query Match:	99.38%	Mismatches:	1						
DB:	14	Indels:	0						
		Gaps:	0						
US-10-082-502-17 (1-182) x BQ421827 (1-880)									
QY	1	MetlsglytTTPGjyTTPleuAlaleuLeuLeuGlyAlaleuLeuGlyThralaTrrpa	20						
Db	102	ATGAAGAGCTGGGGTTCGTCGCCCTCTTCCTTCGGGGCTCTCTGGGAACCCCTGGCT	161						
QY	21	ArgArGserGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspLeuGlnTrp	40						
Db	162	CGGAGGAGCCAGGATCTCCACTCTGGAGATCGACGCGCTCTGGTGAATGAACCTAAGG	221						
QY	41	GluIleuLeuGlnValAspProLysTrpTrlIleGlnMetGlySerPheArgTrlLeuAsnPro	60						
Db	222	GAATTCGCCAGGTGGACCCCAAGAGACCATTCAGATGGGATCTTCGGATCAATCA	281						
QY	61	AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisIleuThrGlu	80						
Db	282	GATGGCAGCCAGTCATGTGGTGGAGGTCCTTATGCCCCGCTCAGAGCCACTCAGAG	341						
QY	81	LeuLeuGlnGluLeuLeuCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr	100						
Db	342	CTCTGGAGAGATATGTGACGGATGAAGAGATATGGGAACAGATTGATCTTCCACC	401						
QY	101	HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu	120						
Db	402	CATCGCAAGAACTACCTAGCTGAATGGGCGCTAATGGAGAAATCAGTGAATCTGGACTTA	461						
QY	121	GlnGlyTlLeuArgTrlLeuAspSerAspLleSerGlyThrLeuLysPheAlaCysGluSerIle	140						

DB 462 CAAGCATCCGATCGACTGACATATTAGCCGACCCTCAACTTTCCGTGTAGACATT 521  
OY 141 VALGUGUUTYrGLuAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnVallys 160  
DB 522 GTGGAGGAAATACGAGGATGACTCATTTGTAATCTTTCCGAGAGGCTGACATGTTANA 581  
OY 161 AspGlySerGlnSerValValGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 60  
DB 582 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 641  
OY 181 GluLeu 182  
DB 642 GAGCTA 647  
RESULT 10  
BI550436 783 bp mRNA linear EST 05-SEP-2001  
LOCUS 60319205F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5264328 5',  
DEFINITION mRNA sequence.  
ACCESSION BI550436  
VERSION BI550436.1 GI:15437748  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 783)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://lmgc.lnl.gov  
Plate: LLM11666 row: 1 column: 01  
High quality sequence stop: 730.  
Location/Qualifiers  
1..783  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5264328"  
/clone\_lib="NIH\_MGC\_95"  
/tissue\_type="hippocampus"  
/lab\_host="DH10B"  
/note="Organ: Brain; Vector: pBluescriptPR (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcaga  
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',  
size-selected for average insert size 2.5 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE COUNT 209 a 189 c 225 g 160 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.77e-110 Length: 783  
Score: 952.00 Matches: 181  
Percent Similarity: 99.45% Conservative: 0  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.17% Indels: 0  
DB: 13 Gaps: 0  
US-10-082-502-17 (1-182) x BI550436 (1-783)  
OY 1 MetGlySerGlnSerValValGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 20

DB 84 ATGAAAGCGTGGGCTGGCTGGCCCTTCCTGGGGGGCCCTGGGAACCCCTGGGCT 143  
OY 21 ArgGlySerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 40  
DB 144 CGGAGGAGCCAGGATCTCCACTGTGGAGCATGCAGGCTCTGTGGATGAACTCAATGG 203  
OY 41 GluIleAlaGluValAspProGlyValThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 204 GAATATGCCACAGGTGAGCCCAAGAAAGACCATTCAGATGGATCTTCCGGATCAATCA 263  
OY 61 AspGlySerGlnSerValValGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80  
DB 264 GATGGCAGCCAGTCACTGAGTGGAGGCTGATATGCCCTGAAGAGCCACCTCACAG 323  
OY 81 LeuLeuGluGluIleCysAspArgMetGlyGlyTrpGlyGluGlnIleAspProSerThr 100  
DB 324 CTGCTGGAGAGATATGTGACCGGATGAAAGAGTATGGGAAACAAATGATCTTCCACC 383  
OY 101 HisArgGlyAsnTrpValArgValAlaGlyArgAsnGlyGluSerGluLeuAspLeu 120  
DB 384 CATGCCAAGAACTACGTAAGCTGTAGTGGCCGGAATGAGAACTGGAACCTGACCTA 443  
OY 122 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuGlyPheAlaCysGluSerIle 140  
DB 444 CAAGCATCCGATCGAATGACTGACATATTAGCCGACCCTCAACTTTCCTGTGAGACATT 503  
OY 141 VALGUGUUTYrGLuAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnVallys 160  
DB 504 GTGGAGGAAATACGAGGATGAACTCATTTGTAATCTTTCCGAGAGGCTGACATGTTANA 563  
OY 161 AspGlySerGlnSerValValGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 180  
DB 564 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 623  
OY 181 GluLeu 182  
DB 624 GAGCTA 629  
RESULT 11  
BQ421069 868 bp mRNA linear EST 23-MAY-2002  
LOCUS BQ421069  
DEFINITION AGENCOURT\_7911956 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6009949  
5', mRNA sequence.  
ACCESSION BQ421069  
VERSION BQ421069.1 GI:21116384  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 868)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: DCTD/DTF/Gazdar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://lmgc.lnl.gov  
Plate: LLM13197 row: d column: 14  
High quality sequence stop: 542.  
Location/Qualifiers  
1..868  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6009949"  
/clone\_lib="NIH\_MGC\_68"  
/tissue\_type="large cell carcinoma"



Db 619 GTGAGAGATAGAGAGACTCAATTCATCTTTCCCGAGAGCGTCACATGTATA 678  
 Oy 161 AspLySerGlnSerValArgThrAspLeuCyAspHisAlaLeuHisSerHisAsp 180  
 Db 679 GACAAACCTTGAGTAGAGACACATCTTTGTGACCATGCGCCCTGCACATTCGCATGAT 738  
 Oy 181 GlnLeu 182  
 Db 739 GAGCTA 744  
 RESULT 13  
 B1550162 773 bp mRNA linear EST 05-SEP-2001  
 LOCUS B1550162  
 DEFINITION 603193045F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5264339 5',  
 mRNA sequence.  
 ACCESSION B1550162  
 VERSION B1550162.1 GI:15437474  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 773)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cga@bbs-rtmail.nih.gov  
 Issue Procurement: Miklos Paikoytis, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: LLM11666 row: 1 column: 12  
 High quality sequence stop: 765.  
 FEATURES  
 Location/Qualifiers  
 source 1..773  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5264339"  
 /clone\_1db="NIH\_MGC\_95"  
 /clone\_2db="NIH\_MGC\_95"  
 /issue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pBluescriptR (modified  
 pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 204 a 178 c 226 g 163 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. NO.: 1.22e-108 Length: 773  
 Score: 939.00 Matches: 181  
 Percent Similarity: 98.91% Conservative: 0  
 Best Local Similarity: 98.91% Mismatches: 1  
 Query Match: 97.81% Indels: 1  
 Gaps: 0  
 US-10-082-502-17 (1-182) x B1550162 (1-773)  
 Oy 1 MetLySerGlnSerValArgThrAspLeuCyAspHisAlaLeuHisSerHisAsp 20  
 Db 84 ATGAAGAGCTGGGGTGGCTGCTGCTGGGGGGCCCTGCGAAGCGCTGGGCT 143  
 Oy 21 ArgArgSerGlnAspLeuHisCysSerValArgAlaLeuValAspGlnLeuGlnUtrp 40

Db 144 CGAGAGCGCAGAGATCTCCACTGTCGAGCATCAGGGCTGTGGATGAACTACAAATGG 203  
 Oy 41 GlnLeaAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60  
 Db 204 GAAATTCGCCAGCTGAGACCCCAAGAACCATTCACATGCGATCTTCCGATCATCA 263  
 Oy 61 AspGlySerGlnSerValValGlnValProTyraAlaArgSerGlnAlaHisLeuThrGlu 80  
 Db 264 GATGCAGCAGCATGCTAGTGTGGAGGTGCTTATGCCCGCTAGAGGCCACCTCACAGAG 323  
 Oy 81 LeuLeuGlnGlnIleCysAspArgMetLysGlnUtrpGlnGlnIleAspProSerThr 100  
 Db 324 CTGCTGAGAGCATATGTGACCGGATGAGAGGTGTGGGAACAATGATCTTCCACAC 383  
 Oy 101 HisArgLySerAsnTyraValArgValAlaGlyArgAsnGlyLysSerGlnLeuAspLeu 120  
 Db 384 CATCCCAAGAACTACGCTACGCTACTGGGCCGAATGGAATCCAGTGAACCTGACCTA 443  
 Oy 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 Db 444 CAAGCATCCGAAATGAGCTACAGATATAGCGGCACCTCAAGTTGCGTGAGAGCAT 503  
 Oy 141 ValGlnGlnUtrpGlnAspGlnLeuLeuIleGlnPheSerArgGlnAlaAspAsnValLys 160  
 Db 504 GTGAGAGATACGAGAGATGAACTCAATGAAATCTTTCCCGAGAGCGTCACAAATGTATA 563  
 Oy 161 AspLySerGlnSerValArgThrAspLeu-CysAspHisAlaLeuHisSerHisAsp 180  
 Db 564 GACAAACCTTGAGTAGAGACACATCTTTGTGACCATGCGCCCTGCACATTCGCATGA 623  
 Oy 180 pGlnLeu 182  
 Db 624 TGAGCTA 630  
 RESULT 14  
 AK013568 754 bp mRNA linear HTC 19-JAN-2002  
 LOCUS AK013568  
 DEFINITION Mus musculus adult male hippocampus cDNA, RIKEN full-length  
 enriched library, clone:2900022B12:transmembrane protein 4, full  
 insert sequence.  
 ACCESSION AK013568  
 VERSION AK013568.1 GI:12850979  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain:C57Bl/6J) adult male hippocampus cDNA to mRNA,  
 clone:2900022B12.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 REFERENCE 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes  
 MEDLINE 20499374  
 PUBMED 11042159  
 REFERENCE 3  
 Shibata, K., Itoh, M., Mizawa, K., Nagaoke, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, U., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,  
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, T., Kita, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format









Db 104 ATGAAGGCTGGGGTGGCTAGCCCTACTTTGGGGTCTGCTGGAACTGCTGGCT 163  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 40  
Db 164 CGAAGGAGCCAAAGATCTACCTGTGAGGCTGTGAGGGCTGTGATGATTAAGTGG 223  
QY 41 GlnIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 224 GAATTTGCCCGGTGGAGCCCAAGAGACCATTCAGATGGATCCCTCCGAATCAATCCA 283  
QY 61 AspGlySerGlnSerValValGlnValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 284 GATGGAGCCACTACGTGTGAGTACTTATGCCGCTCAGAGGCCACCTCACAGAG 343  
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 344 TTGCTTGAGAGGTGTGTGACCGAATGAAGAGTACGGGGAGACAGATTGACCCCTCTAC 403  
QY 101 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGlnSerSerGluLeuAspLeu 120  
Db 404 CACCCGAACAACCTACGTACGCTCGTAGCCGGAATGGGAATCCAGTGAACCTAGACTTAA 463  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 140  
Db 464 CAGGGCATCCGAATGACTCAGATATCAGCGGACACCTCAAGTTGCTGTGAGAGCATT 523  
QY 141 ValGlnGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
Db 524 GTGGAGAAATACGAGGATGAGCTTATCGAATTCCTCTCCAGAGAGGCTGACACGTTAA 583  
QY 161 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisIleSerHisAsp 180  
Db 584 GACAAACTTTGCAGTAACCGAGACAGATCTATGTGACCATGCCCTGACACAGATTCACGAT 643  
QY 181 GluLeu 182  
Db 644 GAGCTA 649

Search completed: March 31, 2003, 04:51:05  
Job time : 1173.85 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1725.3 Seconds  
(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-17  
Perfect score: 960  
Sequence: 1 MKGGMWLLALIGALGTAWA.....LCSKRTDCDHALHISNDEL 182

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spoos/US10082502/runat.24032003.135059.6547/app.query.fasta\_1.1308  
-DB=GenEmbl -QFWT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=tbl -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502@cgn.1.1.4758\_@runat.24032003.135059.6547 -NCPU=6 -ICPU=3  
-NO\_XLPPX -NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_da.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_da.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vl.\*  
30: em\_hg\_hum.\*  
31: em\_hg\_inv.\*  
32: em\_hg\_other.\*  
33: em\_hg\_mus.\*  
34: em\_hg\_pln.\*  
35: em\_hg\_rtd.\*  
36: em\_hg\_mam.\*  
37: em\_hg\_vrt.\*  
38: em\_sy.\*  
39: em\_hggo\_hum.\*  
40: em\_hggo\_mus.\*  
41: em\_hggo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query		Match Length		DB		ID		Description	
No.	Score	%									
1	960	100.0	814	6	AX440456					AX440456 Sequence	
2	960	100.0	814	9	AB015631					AB015631 Homo sapi	
3	960	100.0	1080	9	AY032624					AY032624 Homo sapi	
4	960	100.0	1210	6	AX464040					AX464040 Sequence	
5	938	97.7	798	10	BC008261					BC008261 Mus muscu	
6	938	97.7	1017	10	AF186115					AF186115 Mus muscu	
7	938	97.7	158198	2	AC012013					AC012013 Homo sapi	
8	488.5	50.9	155023	2	AC025574					AC025574 Homo sapi	
9	481	50.1	182350	2	AC073896					AC073896 Homo sapi	
10	481	50.1	184762	2	AC023500					AC023500 Homo sapi	
11	446.5	46.5	207424	10	AC090489					AC090489 Genomic s	
12	436	45.4	101666	2	AC109891					AC109891 Rattus no	
13	369	38.4	649	9	AF186113					AF186113 Homo sapi	
14	369	38.4	824	9	BC001027					BC001027 Homo sapi	
15	328.5	34.2	69753	6	AC103156					AC103156 Rattus no	
16	266	27.7	564	6	AX440375					AX440375 Sequence	
17	246	25.6	2378	6	AK094445					AK094445 Homo sapi	
18	212	22.1	996	3	AY094698					AY094698 Drosophi	
19	190.5	19.8	40824	3	U13070					U13070 Caenorhabe	
20	155.5	16.2	622	5	AB055671					AB055671 Arabidops	
21	154	16.0	785	8	AY088187					AY088187 Drosophi	
22	147	15.3	175963	3	AC009379					AC009379 Drosophi	
23	147	15.3	279530	3	AE003518					AE003518 Drosophi	
24	144	15.0	1286	3	AY060987					AY060987 Drosophi	
25	141	14.7	55359	2	AC012854					AC012854 Drosophi	
26	141	14.7	172904	3	AE003831					AE003831 Drosophi	
27	141	14.7	275390	3	AE003831					AE003831 Drosophi	
28	134.5	14.0	161566	9	AC008060					AC008060 Homo sapi	
29	131.5	13.7	744	6	AR142811					AR142811 Sequence	
30	131.5	13.7	746	6	AX136616					AX136616 Sequence	
31	131.5	13.7	747	6	BD006701					BD006701 Novel pol	
32	131.5	13.7	1413	9	BC019903					BC019903 Homo sapi	
33	131.5	13.7	1419	6	AX136423					AX136423 Sequence	
34	131.5	13.7	1470	6	AX376412					AX376412 Sequence	
35	131.5	13.7	1470	6	AX545716					AX545716 Sequence	
36	131.5	13.7	1470	6	AX491194					AX491194 Sequence	
37	131.5	13.7	1512	6	AR142810					AR142810 Sequence	
38	131.5	13.7	1547	9	BC032339					BC032339 Homo sapi	
39	129.5	13.5	76187	2	AC015219					AC015219 Drosophi	
40	122	12.7	168418	9	AC009403					AC009403 Homo sapi	
41	115	12.0	181896	2	AC114126					AC114126 Rattus no	
42	115	12.0	201649	2	AC120813					AC120813 Rattus no	
43	114.5	11.9	201649	2	AC120813					AC120813 Rattus no	
44	113	11.8	1260	6	AX380553					AX380553 Sequence	
45	111	11.6	210631	2	AC129184					AC129184 Mus muscu	

RESULT 1

ALIGNMENTS

AX440456 AX440456 814 bp DNA linear PAT 28-JUN-2002  
 LOCUS  
 DEFINITION Sequence 309 from Patent WO0190154.  
 ACCESSION AX440456  
 VERSION AX440456.1 GI:21665266  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 AUTHORS 1 Xu,J., Mitcham,J.L., Harlocker,S.L., Dillon,D.C., Secrist,H.,  
 Lodes,M.J., Algate,P.A., Fling,S.P., Mannion,T., Benson,D.R. and  
 Carter,D.  
 TITLE Compositions and methods for the therapy and diagnosis of ovarian  
 cancer  
 PATENT: WO 0190154-A 309 29-NOV-2001;  
 JOURNAL CORIXA CORPORATION (US)  
 FEATURES  
 SOURCE 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.15e-87 Length: 814  
 Score: 960.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 US-10-082-502-17 (1-182) x AX440456 (1-814)  
 Oy 1 Mettysglttprglttprleuallaleuenuleuglualaleuenugttthralatrrala 20  
 |||||||  
 Db 145 ATGAAAGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 Oy 21 Argatserglnaspleuallaleuenuleuglualaleuenugttthralatrrala 40  
 |||||||  
 Db 205 CGAGAGACCCAGATCTCCAGTGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 264  
 Oy 41 Glutleallaglnvalasprpdylysthrllleglmetgllyserphatrglleaspro 60  
 |||||||  
 Db 265 GAAATGGCCAGGTGGACCCCAAGAGACCAATTCAGATGGATCTTCCGATCAATCA 324  
 Oy 61 AspglyserglnserValalgluValaProtyrAlaArgserglualaleuenuhrglu 80  
 |||||||  
 Db 325 GATGGACACGATCGATGAGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAGAG 384  
 Oy 81 Leuleuglulilecysaspargmetylsclutrygluglunlleaspproserthr 100  
 |||||||  
 Db 385 CTGCTGGAGGAGATATGATGACCGGATGAGAGATGAGGAGAACATGATCTCCACAC 444  
 Oy 101 HlsarglysaenttyvalalgluValaIgluArgannglyluserSerGluleuasprou 120  
 |||||||  
 Db 445 CATCCCAAGAACTAGTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGTAC 504  
 Oy 121 GlnglylIleatrglleasSerAsprlleSerGlyThrleuLyPheAlaCysgluSerlle 140  
 |||||||  
 Db 505 CAAGGCATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 564  
 Oy 141 ValgluIutrygluAspgluleuileglupherSerArglualaleuasproulyls 160  
 |||||||  
 Db 565 GTGGAGGAATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 624  
 Oy 161 AspulyseuCyserLyArgThrAspleuCyasprHlsalaleuHlsIleSerHlsasp 180  
 |||||||  
 Db 625 GACAAACTTTCAGTAAACGAGATCTTGTGACATGACCTGCACATATGATGATGAT 684  
 Oy 181 Gluleu 182  
 |||||||

Db 685 GAGCTA 690  
 RESULT 2  
 LOCUS AB015631  
 DEFINITION Homo sapiens mRNA for type II membrane protein, complete cds,  
 clone:HP10390.  
 ACCESSION AB015631  
 VERSION AB015631.1 GI:4586839  
 KEYWORDS type II membrane protein.  
 SOURCE Homo sapiens gastric adenocarcinoma CDNA to mRNA,  
 clone:lib.pKAl-meta-1 clone:HP10390.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 AUTHORS 1 Yokoyama-Kobayashi,M., Yamaguchi,T., Sekine,S. and Kato,S.  
 TITLE Selection of cDNAs encoding putative type II membrane proteins on  
 the cell surface from a human full-length cDNA bank  
 JOURNAL Gene 228 (1-2), 161-167 (1999)  
 MEDLINE 99173880  
 REFERENCE 2 (bases 1 to 814)  
 AUTHORS Kato,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JUN-1998) Setsu Kato, Research Institute of National  
 Rehabilitation Center for the Disabled, Department of  
 Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama  
 359-8555, Japan (E-mail:setsu@rehab.go.jp).  
 Tel:042-995-3100(ex.2568). Fax:042-995-3133)  
 FEATURES  
 SOURCE 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HP10390"  
 /tissue="type II gastric adenocarcinoma"  
 /clone\_lib="pKAl-meta-1"  
 145. 693  
 /product="type II membrane protein"  
 /codon\_start=1  
 /protein\_id="PAA76498.1"  
 /db\_xref="GI:4586840"  
 /translation="MKRGVALLLGALLGTRMKRRSDOLHCGACRLVDELEETIQ  
 VPKKTIOMGSEFRINPDSOSVVEVPAKRSVAHLYLELEICDMKKEGQIDPSTIR  
 KNVRYVVGRRGSESLDQGIKIDSDISDTLKFACESIVEYEDELIEFSREADNVK  
 DKLSKRTDLCDHALHSHDEL"  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.15e-87 Length: 814  
 Score: 960.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 US-10-082-502-17 (1-182) x AB015631 (1-814)  
 Oy 1 Mettysglttprglttprleuallaleuenuleuglualaleuenugttthralatrrala 20  
 |||||||  
 Db 145 ATGAAAGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 Oy 21 Argatserglnaspleuallaleuenuleuglualaleuenugttthralatrrala 40  
 |||||||  
 Db 205 CGAGAGACCCAGATCTCCAGTGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 264  
 Oy 41 Glutleallaglnvalasprpdylysthrllleglmetgllyserphatrglleaspro 60  
 |||||||  
 Db 265 GAAATGGCCAGGTGGACCCCAAGAGACCAATTCAGATGGATCTTCCGATCAATCA 324  
 Oy 61 AspglyserglnserValalgluValaProtyrAlaArgserglualaleuenuhrglu 80  
 |||||||  
 Db 325 GATGGACACGATCGATGAGTGGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAGAG 384



Percent Similarity:	100.00%	Conservative:	
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-082-502-17 (1-182) x AX464040 (1-1210)

[illegible]

Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: amgebcm.tmc.edu  
Gunnaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,  
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, K.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK Plate: 5 Row: b Column: 23  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gl: 9903606.

**FEATURES**  
**Source**

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/map="129, C57BL/6J, FVB/N"
/clone="MGC:6853 IMAGE:2650612"
/tissue_type="Mammary tumor, Brca1-/-; MMTV-Cre model, 10
months old, gross tissue."
/clone_1lb="NCI_CGAP_Mam3"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
115. .663

```

BASE COUNT	
227	<pre> /codon_start=1 /product="transmembrane protein 4" /protein_id="AAH0261.1" /db_xref="GI:14198401" /db_xref="locusid:56530" /translation="MKGWGLALLGLVLLGTAWARRSODLHCACRALVDELEIAR VPRKTIOMGSEFRINPDGQSVDEVPYAREALTELEVCYCRMEYQGIOPSTR KNVRRVSRNGSESELDQIRIDSDISGLTFACESEIYVEYDEDELEFFSREADNYK DKKCSRPTDICALRSHDEL" </pre>
182 c	228 c
161 t	

### Alignment Scores:

Pred. No.:	1,848-85	Length:	758
Score:	938.00	Matches:	1777
Percent Similarity:	98.35%	Conservative:	2
Best Local Similarity:	97.25%	Mismatches:	3
Query Match:	97.71%	Indels:	0
DB:	10	Gaps:	0

US-10-082-502-17 (1-182) x BC008261 (1-798)

QY	1	Melvelylrprpqlrptpleualaleuleuleugllylaleuleuglythrallatpala	20
Db	115	ATGAAGGCTGGGGTGGCTACGCCCTACTTTGGGGGCTCCTCGGGAACTGCTGGGCT	174
QY	21	AtgArGserGlnaspIleuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp	40
Db	175	CGAAGGCGCCAAgTCTACACTGTGGAGCTGTGCAGGGCTCTGTGGTGTGAATTAAGATGG	234
QY	41	GluIleAlaGlnValAspProLysThrIleGluMetGlySerPheArgIleAspPro	60
Db	235	GAATTCCTCCGCGTGGACCCCAAGAGACCATTCAGATGGGATCCTTCCAAATCAATCA	294
QY	61	AspGlySerGlnSerValValGluValProTyrAlaAspSerGluAlaHisLeuThrGlu	80
Db	295	GATGGCAGCAGTCAGTTGTGGAGGTACCTTATGCCCCGTGAGGGCCCACTCACAGAG	354
QY	81	LeuLeuGlnGluIleIcysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr	100
Db	355	TTCTCTTGAGAGGGTGTGTGACCAATTAAGAGATACGGGGAAAGATGACCTTCTACAC	414
QY	101	HisArgLysAsnTyrValArgValIvalGlyArgAsnGlyLysSerGluLeuAspLeu	120
Db	415	CACCGCAAGAACTACGTACCGCTGTGAGCGGGGAATGCAAAATCCAGTAACTACGACTTA	474
QY	121	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuAspHeaLacyGluSerIle	140



```

Db 475 CAGGCATCGCAATTCAGATATCAGCGGACCCCTCAAGTTCCGTGTAGACGAT 534
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 535 GTGGAAGAAATACAGAGATAGCTATTCGAATTTCTCCAGNAGGCTGCACAGTTAA 594
QY 161 AspGluLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
Db 595 GACAAACTTTCAGTAAGCGGACAGATCTATGTGACCATGCTCCGACAGATCTCAGAT 654
QY 181 GluLeu 182
Db 655 GAGCTA 660

RESULT 6
AF186115 1017 bp mRNA linear ROD 13-JAN-2000
LOCUS Mus musculus putative secreted protein Zs1G9 (Zs1G9) mRNA, complete
DEFINITION cds.
ACCESSION AF186115 GI:6014635
VERSION AF186115.1
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1017)
Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and
O'Hara, P.
MUS musculus putative secreted protein
Unpublished
2 (bases 1 to 1017)
Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and
O'Hara, P.
Direct Submission
Submitted (14-SEP-1999) Biomedical Informatics, ZymoGenetics,
Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA
FEATURES
source
1..1017
/organism="Mus musculus"
/db_xref="dbEST:AA049839"
/db_xref="taxon:10090"
/clone="IMAGE:476061"
1..1017
/gene="Zs1G9"
/gene=".891"
343..891
/gene="Zs1G9"
/codon_start=1
/product="putative secreted protein Zs1G9"
/protein_id="AA01433.1"
/db_xref="GI:6014636"
/translation="MKMGWMLALLGLGTAMARRSODLHGACRALYDELEMTAR
VDPKTIOMGSRINPDGSOVVEYARSEAHLETLELDEDRMEKEGEOIDPSNR
KNYRVNRSNGSSSELIDGICRIDSISITLKFACSEIVEYEDELIEFFSRADNVK
DWLCKSRITDCHALHRSDEL"
CDS
gene
/gene="Zs1G9"
343..891
/gene="Zs1G9"
/codon_start=1
/product="putative secreted protein Zs1G9"
/protein_id="AA01433.1"
/db_xref="GI:6014636"
/translation="MKMGWMLALLGLGTAMARRSODLHGACRALYDELEMTAR
VDPKTIOMGSRINPDGSOVVEYARSEAHLETLELDEDRMEKEGEOIDPSNR
KNYRVNRSNGSSSELIDGICRIDSISITLKFACSEIVEYEDELIEFFSRADNVK
DWLCKSRITDCHALHRSDEL"
BASE COUNT 248 a 250 c 315 g 204 t
ORIGIN
Alignment Scores:
Pred. No.: 2,49e-85 Length: 1017
Score: 938.00 Matches: 177
Percent Similarity: 98.35% Conservative: 3
Best Local Similarity: 97.25% Mismatches: 2
Query Match: 97.71% Indels: 0
DB: 10 Gaps: 0
US-10-082-502-17 (1-182) x AF186115 (1-1017)
QY 1 MetLysGlyTyrPglYrPleuAlaLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20
Db 343 ATGAAGGCGGGGTTGGCTAGCCCTACTTTTGGGGGCTCTGCTGGGAATGCTGGGCT 402
QY 21 ATGATSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 40

```

```

Db 403 CGAAGACCCCAAGATCTACAGCTGTGAGCGCTCTGTGATCAATAGAGTCG 462
QY 41 GluIleAlaGluValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 463 GAATTTGCCCGGTGGAGCCCAAGAAAGACCATTCAATGGATGCTTCGAAATCAATCCA 522
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80
Db 523 GATGGCAGCAGTCACTTGTGGAGTACTTATGCGCGCTCAGNAGGCCACCTCCACAGAG 582
QY 81 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 583 TTTCCTTGAGGAGCGTGTGTGACCGAATGAAGAGTACGGGAAACAGATTACCTTCTACC 642
QY 101 HisArgLysAspTyrValArgValValGlyArgAspGlyCysSerSerGluLeuAspLeu 120
Db 643 CACCGCAAGAACTACATACGTCGTGAGCCGGAATGGAATCCAGTGAATCACTTA 702
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 703 CAGGCATCGCAATTCAGTACTAGATATCAGCGGACCCCTCAAGTTTGCCTGTAGACATT 762
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 763 GTGGAAGAAATACAGAGATGAGCTTATCGAATTTCTCTCCAGAGAGCTGCACAGTTAA 822
QY 161 AspGluLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 180
Db 823 GACAAACTTTCAGTAAGCGGACAGATCTATGTGACCATGCTCCGACAGATCTCAGAT 882
QY 181 GluLeu 182
Db 883 GAGCTA 888

RESULT 7
AC012013 158198 bp DNA linear HTG 17-JUL-2001
LOCUS Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT
DEFINITION SEQUENCE, 5 unordered pieces.
ACCESSION AC012013 GI:14717292
VERSION AC012013.18
KEYWORDS HTG; HTGS_PHA5E1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 158198)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Alshrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barberia, J.,
Benton, J., Blinze, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burnett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, R.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinb, H.H., Douthwaite, K.J., Dzipac, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gablitz, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlik, P., Hayes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,
Hollins, B., Homel, F., Howard, S., Huber, J., Huyls, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W.,
Lounseged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawliny, E., McLeod, M.P., Meador, M.,
Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,

```



```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved. 3033: contig of 3033 bp in length
      1

```

BASE COUNT	40008 a	36207 c	35987 g	41693 t	1128 others
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="12" /clone="RP12-348M3"				
Alignment Scores:	2,31e-37				
Pred. No.:	488.50				
Score:	39.728				
Percent Similarity:	39.728				
Best Local Similarity:	39.728				
Query Match:	50.898				
Di:	2				
US-10-082-502-17 (1-182) x AC025574 (1-155023)					
QY 69	ValProtyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluIleCysAspArg				
Db 56120	GTGCTTATGCCCCCGCTCAGAGGCCCCCTCACAGAGCTGCTGGAGGATATGGACCGG				
QY 89	MetIleGluTrpGlyGluGluGlnIleAspProSerThrHisArgIysAsnTrpValArgVal				
Db 56060	ATGAAGAGATATGGGACACGATTGATCTTCACCCATCCCAAGAACTCGACAGCTGTGA				
QY 109	ValGlyArgAsnGlyLeuSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp				
Db 56000	GTGGGCGGGAATGGAACATCCAGTGAACAGCAGCTACAAAGCATCGAATCGACTCAGAT				
QY 129	IleSerGlyThrLeuLysPheLeu-----				
Db 55940	ATTAGCGGCGCACCTCTCAAGTTTCGCTGAGCTATGGGAATCGGTAGCTGTTCTTGGACAT				
QY 136	-----				
Db 55880	TAAGGGGTGTGGAGAACCAATTAGAAAGTTAGGCTGATATCTTGTCTCTCTTTGG				
QY 136	-----				
Db 55820	AAGTGAAGCAGAGAGCCCTTATTGGCCCTGTGCACCCATTCTCCCTTGGTTGGCAG				
QY 137	CysGluSerIleValGluGluTrpGluAspIleuLeuIleGluPhePheSerArgGluIle				
Db 55760	TGTGAAGCATTTGTGGAGGAATACAGAGATGACTATTAATTTCTTCCGAGAGGCT				
QY 157	AspAsnValIleAspLysLeuCysSerIleArgThr-----				
Db 55700	GCAACATCTTAAGACAAACTTTGGACGATTAAGGAACAGTAAAGCTGACCCCACTTATATCTC				

168 ----- 168  
Db 55640 CTGTTCTAGTACCTCTGTGGAGCTGACATTCCTTACTACACCATACACCCCATCCCA 55581  
168 ----- 168  
Db 55580 GAAACCTCTGATTAGGTCAAGACCTGCTTCCATTTCCCTTGCTTACGGGACTTGGG 55521  
168 ----- 168  
Db 55520 GTGTTCTCAGCATTTTCTGTCAGCTTCAGATTTGTGCTCCCTTGACCTGAGATACA 55461  
168 ----- 168  
Db 55460 GGTGAGAGACAGTGGGAGTGTATTATCATTTATAGCATATATTCTGACAGATGTGGT 55401  
168 ----- 168  
Db 55400 TGGTGTGTTTGTCTGACGCTGCTGCGAATTTCTTCTGATCTGACGATATACCTA 55341  
169 -----  
Db 55340 TGAGAAATGACATATTCCTTAATGTTTGGCATTGGACATCTTGTGACATGCTG 55281  
176 HISTSERHISAPGILen 182  
|||||  
Db 55280 CACATATGACATGATGAGCTA 55260  
RESULT 9  
AC073896/c 182350 bp DNA linear HTG 24-AUG-2002  
LOCUS Homo sapiens chromosome 12 clone RP11-977G19, WORKING DRAFT  
DEFINITION AC073896  
AC073896.37 GI:22450359  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULFILL; HTGS\_ACTIVEPIN.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 182350)  
Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C.,  
Alshrocks,S.L., Amaralungu,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonan,D.,  
Bouck,J., Bowler,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bulay,C., Burch,P., Burkelt,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Dayle,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dihn,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorell,D.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,R., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homsli,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichteberg,O., Liew,C., Liu,J., Liu,M., Lounsged,H.,  
Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Matlindala,A., Martinez,E.,  
Massey,E., Maswiny,E., McLeod,M.P., Meador,M., Mel,C., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabdat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newson,J., Newton,N., Nguyen,N., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogund,M., Okunodu,G.,  
Oragunye,N., Oyler,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,

Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Sytek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S.,  
Umanal,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,O.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleciyk,R., Woodson,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.  
Unpublished  
Direct Submission  
2 (bases 1 to 182350)  
Worley,K.C.  
Submitted (01-JUL-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 182350)  
Worley,K.C.  
Direct Submission  
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Aug 23, 2002 this sequence version replaced gi:22296903.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Drafting Center Code: BCM  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: HBW1  
Center clone name: RP11-977G19  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Sequencing vector: M13;  
Chemistry: Dye-primer Bodypy: 15% of reads  
Chemistry: Dye-terminator Big Dye: 85% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 179279 bases at least 40  
Consensus quality: 180581 bases at least Q30  
Consensus quality: 181073 bases at least Q20  
Estimated insert size: 17425; sum-of-coverage estimation  
Quality coverage: 8x in Q20 bases; sum-of-coverage estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank.draft.data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 2893: contig of 2893 bp in length  
2894 2993: gap of unknown length  
2994 6686: contig of 3693 bp in length  
6687 6787: gap of unknown length  
6787 17251: contig of 10465 bp in length  
17251 17352: gap of unknown length  
17352 85894: contig of 68543 bp in length  
85894 85995: gap of unknown length  
85995 182350: contig of 96356 bp in length.  
Location/Qualifiers  
1. 182350  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-977G19"

BASE COUNT 46453 a 45416 c 43855 g 46174 t 452 others  
ORIGIN  
Alignment Scores:

```

Pred. No.: 1.61e-36 Length: 182350
Score: 481.00 Matches: 103
Percent Similarity: 63.80% Conservative: 1
Best Local Similarity: 63.19% Mismatches: 6
Query Match: 50.10% Indels: 53
DB: 2 Gaps: 1
US-10-082-502-17 (1-182) x AC07896 (1-182350)

QY 69 ValProTyAlaIarSergIuaIaHIsLeuThgluIeurengluIueCysAspArg 88
|||||
Db 172027 GTCCTTAATGCCCCGTCACAGAGCCACCTCACAGACTGCTGAGAGATATGACCG 171968
|||||
QY 89 MetLysgluTygluIugluInIleasProSerThrHisArgLysanTyValArgval 108
|||||
Db 171967 ATGAGAGATATGCGGAGACAGATGATCTCTCCACCATCGCAAGACATGACGTAGTA 171908
|||||
QY 109 ValGlyArgAsnGlyLysSerSergIuleAspLeuInglyIleArgIleAspSerAsp 128
|||||
Db 171907 GTGGCCGGAAGAGAAATCCACTGACATGACCTACAAAGCATCCGAATCGACTGAGAT 171848
|||||
QY 129 IleSerGlyThrLeuLysPheAla----- 136
|||||
Db 171847 ATTAGCGGCACCTCAAGTTGGCGTGACGTATGGAATCGTACGTCTTTCGACAT 171788
|||||
QY 136 ----- 136
Db 171787 TAAGGGTTGTGGAGAACCATTAAGAAAGTAGCTGATCTTGTCTCTCTTTGG 171728
|||||
QY 136 ----- 136
Db 171727 AAGTGAAGCGAAGAGCCCTTATATGCCCCCTGTACACCCATTTCTCCCTGGGTGGCAG 171668
|||||
QY 137 CysGluSerIleValGluGlyArgIuAspGluLeuIleGluPhePheSerArgIuAla 156
|||||
Db 171667 TGTGACAGCATTTGTGAGAGATACGAGAGATCACTCAATCTTTTCCAGAGACCT 171608
|||||
QY 157 AspAsnValLysAspLysLeuCysSerLysArgThrAsp-LeuCysAspHisAlaLeuH1 176
|||||
Db 171607 GACAAATGTTAAAGCAAACTTTGACATAGCGAAGTAGTAACTGCCCCACTTTATCTC 171548
|||||
QY 176 IleSer 178
|||||
Db 171547 CTCTTCT 171541

RESULT 10
AC023500 184762 bp DNA 1linear HMG 24-AUG-2002
LOCUS
DEFINITION
AC023500 Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT
SEQUENCE
AC023500.32 GI:22095070
VERSION
AC023500 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 184762)
Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alshrocks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbista,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,U., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Burch,P., Burckell,C., Butrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,R., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoggess,M., Holloway,C., Hollins,B.,
Hosni,F., Howard,S., Huber,J., Huij,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,
Karlsom,E., Kelly,S., Khan,U., King,L., Kovach,J., Kovar,C.,
Kraovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Merindale,A., Martinez,E.,
Massey,E., Mayhew,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,K., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokoko,S., Ogih,M., Okunogu,G.,
Oragunde,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pul,L., Quiles,M., Ren,Y.,
Rivera,M., Rojas,A., Rojiboken,I., Rolle,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shoshart,N., Slason,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,C.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 184762)
Worley,K.C.
Direct Submission
Submitted (15-FEB-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 184762)
Worley,K.C.
Direct Submission
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Aug 4, 2002 this sequence version replaced gi:22094227.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAPC
Center clone name: RP11-183H16
----- Summary Statistics
Sequencing vector: Plasmid;
Sequencing vector: M13;
Chemistry: Dye-primer; Body: 2% of reads
Chemistry: Dye-terminator; Big Dye; 98% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 183845 bases at least Q40
Consensus quality: 184849 bases at least Q30
Consensus quality: 185447 bases at least Q20
Estimated insert size: 159026; sum-of-coverage estimation
Quality coverage: 8x in Q20 bases; sum-of-coverage estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 9 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 2040: contig of 2040 bp in length
* 2041 2140: gap of unknown length

```

```
*      2141      4170: contig of 2030 bp in length
*      4171      4270: gap of unknown length
*      4271      8294: contig of 4024 bp in length
*      8295      8394: gap of unknown length
*      8395      14196: contig of 5802 bp in length
*      14197      14296: gap of unknown length
*      14297      33076: contig of 18780 bp in length
*      33077      33176: gap of unknown length
*      33177      52295: contig of 19119 bp in length
*      52296      52395: gap of unknown length
*      52396      84185: contig of 31790 bp in length
*      84186      84285: gap of unknown length
*      84286      120023: contig of 35738 bp in length
*      120024      120123: gap of unknown length
*      120124      184762: contig of 64639 bp in length.
*
FEATURES
  source
    1..184762
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="12"
      /clone="RP11-183h16"

BASE COUNT   47425 a 44751 c 45101 g 46673 t      812 others
ORIGIN
Alignment Scores:
Pred. No.:      1.63e-36      Length:      184762
Score:          481.00      Matches:      103
Percent Similarity: 63.80%      Conservative: 1
Best Local Similarity: 63.19%      Mismatches: 6
Query Match:    50.10%      Indels:      53
DB:             2      Gaps:      1

US-10-082-502-17 (1-182) x AC093500 (1-184762)
OY      69  ValProYrAlArSergLUAhIstLeuthrgLueuLeuGluGluIleCysAspArg 88
         |||||||
DB 43486 GTGCGTTATGCCCGCTCAGAGGCCCACTCACAGAGCTGCTGGAGAGATATGTACCGG 43545
         |||||||
OY      89  MetLysGluYrGluGluGlnIleAspProSerThrHISatgLYsAsnTYrValArgVal 108
         |||||||
DB 43546 ATGAAGGAGTATGGGGAACAATTCATCTTCACCCATCCAGAACTACAGTACGTGA 43605
         |||||||
OY      109 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
         |||||||
DB 43606 GTGGCCGGAATGGAGATCCAGTCACTGACCTCAAGGCATCCGAATCGACACAGAT 43665
         |||||||
OY      129 IleSerGlyThrLeuLYsPheAla----- 136
         |||||||
DB 43666 ATTAGCGGCACCTCAAGTTCGGCGTGAGCTATGGGAATCGGTACGTCTTGACAT 43725
         |||||||
OY      136 ----- 136
         |||||||
DB 43726 TAAAGGATTGTGGNAACAATTAGAAAGTATAGCTGATCTTGTCTCTCTTTGG 43785
         |||||||
OY      136 ----- 136
         |||||||
DB 43786 AAGTGAAGGCAAGAGCCTTATATGCCCCGTGTGCACCCATTCTCCCTGGGTGGCAG 43845
         |||||||
OY      137 CysGluSerIleValGluGluTYrGluAspLueuIleGluPheSerArgGluAla 156
         |||||||
DB 43846 TGTGAGACATTTGGAGAGATACGAGATGACATCAATCTTTTCCCGAGAGCT 43905
         |||||||
OY      157 AspaAsnValLYsAspLYsLeuCYsSerLYsArgThrAsp-LeuCYsAspHISAlaLeuH1 176
         |||||||
DB 43906 GACATATGTTAAAGCAAACTTTCAGATNAGCAACAGTAAAGTCCGCCACTTATATCTC 43965
         |||||||
OY      176 sIleSer 178
         |||
DB 43966 CTGTTCT 43972
         |||

RESULT 11
AC090489
LOCUS      AC090489      207424 bp      DNA      linear      ROD 26-JAN-2002
```

```
DEFINITION      Genomic sequence for Mus musculus, clone RP23-104010, complete
sequence.
ACCESSION      AC090489
VERSION        AC090489.8  GI:18376849
KEYWORDS
SOURCE
  ORGANISM
    Mus musculus.
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  AUTHORS
    McCombie, W.R., de la Bastide, M., Spiegel, L., Preston, R.,
    Kirchhoff, K., Kuhl, K., Nascimento, L., Zuber, T., Bailly, V.,
    Bell, M., Baker, J., Santos, L., Miller, B., Katzenberger, F.,
    Muller, S., King, L., Yang, C., Palmer, L., O'Shaughnessy, A. and
    Dedhia, N.
  TITLE
    Genomic sequence for Mus musculus, clone RP23-104010, complete
    sequence
  JOURNAL
    Unpublished
  REFERENCE
    2 (bases 1 to 207424)
  AUTHORS
    McCombie, W.R.
  TITLE
    Direct Submission
  JOURNAL
    Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Sequencing
    Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
    Harbor, NY 11724, USA
  REFERENCE
    3 (bases 1 to 207424)
  AUTHORS
    McCombie, W.R.
  TITLE
    Direct Submission
  JOURNAL
    Submitted (26-JAN-2002) Lita Annenberg Hazen Genome Sequencing
    Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
    Harbor, NY 11724, USA
  COMMENT
    On Jan 26, 2002 this sequence version replaced gi:18201765.
    This sequence was finished as follows unless otherwise noted: all
    regions were either double-stranded or sequenced with an alternate
    chemistry or covered by high quality data (i.e., phred quality >=
    30); an attempt was made to resolve all sequencing problems, such
    as compressions and repeats; all regions were covered by at least
    one plasmid subclone or more than one M13 subclone; and the
    assembly was confirmed by restriction digest.
  FEATURES
    source
      1..207424
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /clone="RP23-104010"
        /clone_11b="RPC1-23"
        48132..48165
          misc-feature
            /note="We believe the assembly to be correct. The sequence
            is a mononucleotide (T) repeat in which the exact number
            of Ts is unknown. Other subclones in the region show one
            fewer T than that represented in the assembly."
    BASE COUNT   52613 a 48671 c 49663 g 56477 t
    ORIGIN
Alignment Scores:
Pred. No.:      5.61e-33      Length:      207424
Score:          446.50      Matches:      98
Percent Similarity: 48.77%      Conservative: 1
Best Local Similarity: 48.28%      Mismatches: 1
Query Match:    46.51%      Indels:      103
DB:             10      Gaps:      1

US-10-082-502-17 (1-182) x AC090489 (1-207424)
OY      69  ValProYrAlArSergLUAhIstLeuthrgLueuLeuGluGluIleCysAspArg 88
         |||||||
DB 75789 GTACCTTATGCCCGCTCAGAGGCCCACTCACAGAGTTGCTTGAAGAGTGTGTACCGA 75848
         |||||||
OY      89  MetLysGluYrGluGluGlnIleAspProSerThrHISatgLYsAsnTYrValArgVal 108
         |||||||
DB 75849 ATGAAGGAGTACGGGGAACAGATTGACCTTCAACCCACGAGAACTACGTACGCGTC 75908
         |||||||
OY      109 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
         |||||||
DB 75909 GTGACCGGGAATGGAGATCCAGTCACTTACAGGCGATCCGAATTGACTACAGAT 75968
         |||||||
```



OY 129 IIESeGlyThLeulysPheala----- 136  
 DB 75969 ATCAGGGGACCCCTCAAGTTGGGGAGTACCGCTGCGGTAGTAGCTGTTCTTG 76028  
 OY 136 ----- 136  
 DB 76029 GGAGTTATAGGATWCCCTCGGACAGGATGGAGATGATGCGGAAGTGTGCTTAC 76088  
 OY 136 ----- 136  
 DB 76089 TTGTAACTTAAGCAGCTCCAGGGGAGCGGTGGAGATCAGCTTAAGTTGAGCGTGGC 76148  
 OY 136 ----- 136  
 DB 76149 CTAGTTCATAGTAGTTCGAGGCGGAGCTGGGACAGATATAGTAGGCCCTGTCTATAAAT 76208  
 OY 136 ----- 136  
 DB 76209 CCCACCCCCCACAACCAAGTAGAGAGCTGAGCCAGCTGTTGCCAGCTTGGGA 76268  
 OY 137 ----- 145  
 DB 76269 TCTGGGCTAATTTCTGCTCTTCTTGGTCCGAGCTGTGAGAGCATTTGTGGAAGATACGAG 76328  
 OY 146 ASPGLuLeuIleGluPhePheSerArgLualaAspAsnValLysAspLysLeuCySer 165  
 DB 76329 GATGAGCTATACGAATTTCTCTCCAGAGAGGCGTACAGACTTAAGACAACTTTGAGT 76388  
 OY 166 LysArgThr 168  
 DB 76389 AAGCGGACA 76397

RESULT 12  
 AC109891  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-31966, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 47 unordered pieces.

ACCESSION AC109891  
 VERSION AC109891.3 GI:21738196  
 KEYWORDS HTG: HTGS.PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

## REFERENCE

AUTHORS

1 (bases 1 to 101666)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C., Alshrocks,S.L., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T., Barbarella,J., Benton,J., Bimagne,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowles,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denu,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homel,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kravtsov,C., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtenarge,O., Lieu,C., Liu,J., Liu,W., Louisaged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mavhiney,E., McLeod,T.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

## TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nokenko,S., Ogun,M., Okunolu,G.,  
 Orahunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Plickens,R., Primus,E., Pu,L.L., Quiles,M., Remy,Y.,  
 Rives,M., Rojas,A., Rojuben,I., Ruffe,M., Ruiz,S., Savary,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshatari,N., Sisson,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczek,R., Woodem,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.  
 Unpublished  
 2 (bases 1 to 101666)  
 Worley,K.C.  
 Direct Submission  
 Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 101666)  
 Worley,K.C.  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced gi:18860222.

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 Project information  
 Center project name: GSKY  
 Center clone name: CH230-31966  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 60229 bases at least Q40  
 Consensus quality: 63959 bases at least Q30  
 Consensus quality: 66730 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 47 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1  
 1423: contig of 1423 bp in length  
 1523: gap of unknown length  
 1524: contig of 1449 bp in length  
 2973: gap of unknown length  
 3072: gap of unknown length  
 3073: contig of 1645 bp in length  
 4717: gap of unknown length  
 4818: gap of unknown length  
 6632: contig of 1815 bp in length  
 6732: gap of unknown length  
 6733: gap of unknown length  
 8040: contig of 1308 bp in length  
 8041: gap of unknown length  
 8141: gap of unknown length  
 10043: contig of 1903 bp in length  
 10044: gap of unknown length  
 10144: contig of 1094 bp in length  
 11237: gap of unknown length  
 11338: contig of 1025 bp in length  
 11362: gap of unknown length  
 12462: gap of unknown length  
 12463: gap of unknown length  
 14100: contig of 1638 bp in length  
 14101: gap of unknown length



14201 15840: contig of 1640 bp in length  
15841 15940: gap of unknown length  
15941 17053: contig of 1113 bp in length  
17054 17153: gap of unknown length  
17154 18836: contig of 1683 bp in length  
18837 18936: gap of unknown length  
18937 20137: contig of 1201 bp in length  
20138 20237: gap of unknown length  
20238 21736: contig of 149 bp in length  
21737 21837: gap of unknown length  
21837 23619: contig of 1783 bp in length  
23620 23719: gap of unknown length  
23720 25907: contig of 2188 bp in length  
25908 26007: gap of unknown length  
26008 27406: contig of 139 bp in length  
27407 27506: gap of unknown length  
27507 28837: contig of 1331 bp in length  
28838 28937: gap of unknown length  
28938 30229: contig of 1292 bp in length  
30230 30329: gap of unknown length  
30330 32012: contig of 1683 bp in length  
32013 32112: gap of unknown length  
32113 33198: contig of 1086 bp in length  
33199 33298: gap of unknown length  
33299 35519: contig of 2221 bp in length  
35520 35619: gap of unknown length  
35620 37677: contig of 2058 bp in length  
37678 37777: gap of unknown length  
37778 38936: contig of 1159 bp in length  
38937 39036: gap of unknown length  
39037 40273: contig of 1237 bp in length  
40274 40373: gap of unknown length  
40374 41906: contig of 1533 bp in length  
41907 42006: gap of unknown length  
42007 44506: contig of 2500 bp in length  
44507 44606: gap of unknown length  
44608 46803: contig of 2197 bp in length  
46804 46903: gap of unknown length  
46904 48226: contig of 2323 bp in length  
48227 49326: gap of unknown length  
49327 51433: contig of 2107 bp in length  
51434 51533: gap of unknown length  
51534 53763: contig of 2230 bp in length  
53764 53863: gap of unknown length  
53864 56133: contig of 2750 bp in length  
56134 56713: gap of unknown length  
56714 58873: contig of 2060 bp in length  
58874 61527: contig of 2654 bp in length  
61528 61627: gap of unknown length  
61628 63192: contig of 1565 bp in length  
63193 63292: gap of unknown length  
63293 65362: contig of 2070 bp in length  
65363 65462: gap of unknown length  
65463 67438: contig of 1976 bp in length  
67439 67538: gap of unknown length  
67539 70505: contig of 2967 bp in length  
70506 70605: gap of unknown length  
70606 74524: contig of 3919 bp in length  
74525 74624: gap of unknown length  
74625 76737: contig of 2113 bp in length  
76738 76837: gap of unknown length  
76839 80133: contig of 3296 bp in length  
80134 80233: gap of unknown length  
80234 83122: contig of 2889 bp in length  
83123 83222: gap of unknown length  
83223 86731: contig of 3509 bp in length  
86732 86831: gap of unknown length  
86832 89925: contig of 3098 bp in length  
89930 90029: gap of unknown length  
90030 93107: contig of 3078 bp in length  
93108 93207: gap of unknown length  
93208 97419: contig of 4212 bp in length

97420 97519: gap of unknown length  
97520 101666: contig of 4147 bp in length.  
101666  
1. 101666  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-319B6"  
BASE COUNT 24650 a 22934 c 23569 g 23059 t 7454 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 2 64e-32 Length: 101666  
Score: 436.00 Matches: 99  
Percent Similarity: 49.75% Conservative: 0  
Best Local Similarity: 49.75% Mismatches: 1  
Query Match: 45.42% Indels: 99  
DB: 2 Gaps: 1  
US-10-082-502-17 (1-182) x AC109891 (1-101666)  
QY 69 VALPOTYRALAARGSERGIUNLAHLSLEUTHRGILEUleuGIUleuIleCYASPARg 88  
|||||  
DB 97572 GTTCCTTAATGCCCGCTCAGAGGCCACCTCAGAGAGTTGCTTGAGAGATATGACCGA 97631  
|||||  
QY 89 MetLysGIUtyrGIyGIUgInIleAspProSerThrHLSArgLYsAntIyValArgVal 108  
|||||  
DB 97632 ATGAAGGAGTATGGGAGACAGATTGACCTTACCCACCCAGAAAGTACAGTCGCTGT 97691  
|||||  
QY 109 VALGIARGASNGILGUSERSERGIleuAspIleuAspIleuArgIleAspSerAsp 128  
|||||  
DB 97692 GTGGCCGAGAACGGAGGATCCAGTACACTAGACTTACAGGATATCGAATGATTCAGAT 97751  
|||||  
QY 129 ILeSerGIyThrIleuLYsPhe----- 135  
|||||  
DB 97752 ATCAGTGCACACCTCAAGTTGCGGTGAGTATGCTGCGGTGAGTACAGTGTCTTG 97811  
|||||  
QY 135 ----- 135  
DB 97812 GGAGTTCAATAGGAAVTCTGGGCACAGATTGGAGTCGGCGCGGATATGTGTTCATG 97871  
|||||  
QY 135 ----- 135  
DB 97872 CTTGTAAACCGCGGCGAGTGGAGAGCAGTGTGGAAGATTGCACTACTACTGAGT 97931  
|||||  
QY 135 ----- 135  
DB 97932 TCTAGCCAGCTGAGCGATATCATPAAAAACCTGTCTATAAAATCCACAGTCCACCAA 97991  
|||||  
QY 135 ----- 135  
DB 97992 AAACGAGGAGAGAGCTGAGCCGGTGTCTGCCCGCGCTGAGATCCTGGGTACTTCT 98051  
|||||  
QY 136 -----Ala-CysIleSerIleValGIUgInIyrgIuAspGIUleuIleGI 150  
|||||  
DB 98052 CTGTCTTCTGTGCGCAGTGTGAGAGCATTTGTGAGAGAGTTCGAGATTAATTGA 98111  
|||||  
QY 150 uPhePheSerArgIuAlaAspAsnValIyAspIyLeuCYsSerIySArgTrp 168  
|||||  
DB 98112 ATTCTTTCAAGAGAGCGTCAACGTTAAAGACAACTTGCAGTAAAGCGGACA 98166  
|||||  
RESULT 13  
AF186113 649 bp mRNA linear PRI 13-JAN-2000  
LOCUS Homo sapiens putative secreted protein ZS109 (ZS109) mRNA, complete  
DEFINITION cds.  
ACCESSION AF186113  
VERSION AF186113.1 GI:6014631  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE Mammalia: Eutheria: Chordata: Craniata: Vertebrata: Euteleostomi;  
1 (bases 1 to 649) Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.

AUTHORS Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehnert, J. and O'Hara, P.

TITLE Homo sapiens putative secreted protein

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 649)

AUTHORS Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehnert, J. and O'Hara, P.

TITLE Direct Submission

JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave., East, Seattle, WA 98102, USA

FEATURES

source

1..649

/organism="Homo sapiens"

/db\_xref="dbest:R15306"

/db\_xref="taxon:9606"

/chromosome="12"

/map="12q15"

/clone="IMAGE:29675"

1..649

/gene="ZS1G9"

104..358

/gene="ZS1G9"

/codon\_start=1

/product="putative secreted protein ZS1G9"

/protein\_id="AA01431.1"

/db\_xref="GI:6014632"

/translation="MKGMWALILGALLGTANARRSODLHGACRALVDELEMEIAQVDPKRTIOMSGFRINDGSQSVVEVTVPNNKVAHSFG"

BASE COUNT 175 a 150 c 171 g 149 t 4 others

ORIGIN

Alignment Scores:

Pred. No.: 2,76e-28 Length: 649

Score: 369.00 Matches: 69

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 38.44% Indels: 0

Gaps: 0

US-10-082-502-17 (1-182) x AF186113 (1-649)

OY 1 MetlysglyttrpglyttrpaualeuLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20

DB 104 ATGAAGGCTGGGGTGGCTGGCCCTCTCTGGGGGCCCTGGGAGACCGCTGGGCT 163

OY 21 ArgArsgserGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40

DB 164 CGAGGAGCCAGGATCTCCACTGTGGAGCATGCGAGGCTCTGTGGATGAACTAGATGG 223

OY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60

DB 224 GAAATGGCCAGGCTGGAGCCCAAGAACCATTCAGATGGAGATCTTCCGATCAATCCA 283

OY 61 AspGlySerGlnSerValValGluVal 69

DB 284 GATGGCAGCCAGTCAGTGTGGAGGTA 310

RESULT 14

BC001027

LOCUS BC001027 824 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, transmembrane protein 4, clone MGC:1545

IMAGE:3344786, mRNA, complete cds.

ACCESSION BC001027.1 GI:12654402

VERSION BC001027.1

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 824)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (17-NOV-2000) National Institutes of Health, Mammalian

REMARK

COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILN)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)

Gaithersburg, Maryland.

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgcenhgr1.nih.gov](mailto:nisc_mgcenhgr1.nih.gov)

Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McElroy, J.C., McDowell, J., Pearson, R., Snyder, B., Stantrop, S., Thomas, P.J., Tlonson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>

Series: IRAL Plate: 5 Row: P Column: 13

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6014631.

FEATURES

source

1..824

/organism="Homo sapiens"

/db\_xref="locusid:10330"

/db\_xref="taxon:9606"

/clone="MGC:1545 IMAGE:3344786"

/rissue\_type="Eye, retinoblastoma"

/clone\_lib="NIH\_MGC\_16"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

250..504

/codon\_start=1

/product="transmembrane protein 4"

/protein\_id="AAH01027.1"

/db\_xref="GI:12654403"

/translation="MKGMWALILGALLGTANARRSODLHGACRALVDELEMEIAQVDPKRTIOMSGFRINDGSQSVVEVTVPNNKVAHSFG"

BASE COUNT 221 a 194 c 233 g 176 t

ORIGIN

Alignment Scores:

Pred. No.: 3,71e-28 Length: 824

Score: 369.00 Matches: 69

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 38.44% Indels: 0

Gaps: 0

US-10-082-502-17 (1-182) x BC001027 (1-824)

OY 1 MetlysglyttrpglyttrpaualeuLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTrpAla 20

DB 250 ATGAAGGCTGGGGTGGCTGGCCCTCTCTGGGGGCCCTGGGAGACCGCTGGGCT 309

OY 21 ArgArsgserGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40

DB 310 CGAGGAGCCAGGATCTCCACTGTGGAGCATGCGAGGCTCTGTGGATGAACTAGATGG 369

OY 41 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60

DB 370 GAAATGGCCAGGCTGGAGCCCAAGAACCATTCAGATGGAGATCTTCCGATCAATCCA 429

OY 61 AspGlySerGlnSerValValGluVal 69

DB 430 GATGGCAGCCAGTCAGTGTGGAGGTA 456



Search completed: March 31, 2003, 03:37:06  
Job time : 1882.3 secs

*	37672	37771:	gap of unknown length
*	37772	39476:	contig of 1705 bp in length
*	39477	39576:	gap of unknown length
*	39577	41221:	contig of 1645 bp in length
*	41222	41321:	gap of unknown length
*	41322	41530:	contig of 1809 bp in length
*	41311	43230:	gap of unknown length
*	43231	44734:	contig of 1504 bp in length
*	44735	44834:	gap of unknown length
*	44835	46199:	contig of 1365 bp in length
*	46200	46299:	gap of unknown length
*	46300	49061:	contig of 2762 bp in length
*	49062	49161:	gap of unknown length
*	49162	51059:	contig of 1898 bp in length
*	51060	51159:	gap of unknown length
*	51160	53057:	contig of 1908 bp in length
*	53068	53167:	gap of unknown length
*	53168	54641:	contig of 1474 bp in length
*	54642	54741:	gap of unknown length
*	54742	57496:	contig of 2755 bp in length
*	57497	57596:	gap of unknown length
*	57597	59003:	contig of 1407 bp in length
*	59004	59103:	gap of unknown length
*	59104	60621:	contig of 1518 bp in length
*	60622	60721:	gap of unknown length
*	60722	62709:	contig of 1988 bp in length
*	62710	62809:	gap of unknown length
*	62810	64884:	contig of 2075 bp in length
*	64885	64984:	gap of unknown length
*	67553	contig of 2659 bp in length	
*	67554	67753:	gap of unknown length
*	67754	69753:	contig of 200 bp in length
FEATURES	Location/Overlap		

BASE COUNT	ORIGIN	15260 a	17000 c	17392 g	15839 t	4262 others
15260	17000	17392	15839	4262	others	

Alignment Scores:	
Pred. No.:	1,1e-21
Score:	328.50
Percent Similarity:	56.67%
Best Local Similarity:	55.83%
Query Match:	34.22%
DB:	2
	1
	69753
Length:	67
Matches:	
Conservative:	1
Mismatches:	3
Indels:	49
Gaps:	1

US-10-082-502-17 (1-182) \* AC103156 (1-69753)

QY	1	MeltyGGLyTPbGLyTPbLLeuValLeuLeuGlyValLeuLeuGlyLThrAlaTrpAla	20
Db	15128	ATGAAAGGCTGGGGTGGCTGGCCCTACATTCTTGGGGTCTCTGGGAACCTACCTGGCTT	15068
QY	21	ArgArgSerGlnAspLeuHIsCysGly-----	29
Db	15068	CGAAGGAGCCAGGATCTACACTGTGGAGTAAAGCACAATATACAGTGAAGAGGGA	15009
QY	29	-----	29
Db	15008	GGNAGCCGACCCCTGGGAGCTATTGATGCCAAGCTTGGACAAAGATGATGAAGCC	14949
QY	30	-----AlaCys	31
Db	14948	TTGGGTTGACCCATCCCTCCAGATCCCCACATCCCTATTCTCTGGCCGACCTTGC	14889
QY	32	ArgAlaLeuValAspGluLeuGluLTrpGluLeuAlaGlnValAspProLysLThrIle	51
Db	14888	AGGGCTCTGGTGGATGAATTACAGTGGGAATATGGCCCGGGACCCCAAGAAGACATT	14829
QY	52	GlnMetGlySerPheArgILeasnProAspGlySerGlnSerValGluValProLys	71
Db	14838	CAGATGGGATCCCTCCGAATCAATCCAGATGGCAGCCAGCTAGTTGTGGAGTAACTTAC	14769



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 30, 2003, 18:34:08 ; Search time 2881.45 Seconds

Search time 2881.45 Seconds  
(without alignments)  
10796.945 Million cell updates/sec

Title:	US-10-082-502-18
Perfect score:	1069

Perfect score: 1069  
Sequence: 1 gaattcgacaggggggt.....aaaaaaaaattccaccaccac 1069

Scoring table: IDENTITY\_NUC

Searched: 2054640 seqs, 14551402878 residues

number of hits satisfying chosen parameters: 4109280

Maximum DB seq length: 20000000000

```
Post-processing: Minimum Match 08
                  Maximum Match 1008
```

Database :

1:	gb_ba:*
2:	gb_hn:*
3:	gb_in:*
4:	gb_cm:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sts:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_v1:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_v1:*
30:	em_htg_hum:*
31:	em_htg_in:*
32:	em_htg_other:*
33:	em_htg_mus:*
34:	em_htg_pln:*
35:	em_htg_rod:*
36:	em_htg_mam:*
37:	em_hgt_vyt:*
38:	em_sy:*
39:	em_htgo_hum:*
40:	em_htgo_mus:*
41:	em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

C	No.	Score	Query		DB	ID	Description
			Match	Length			
C 1	1	1017	95.1	1017	10	AF186115	AF186115 Mus musc
C 2	2	795.2	74.4	158198	2	AC012013	AC012013 Homo sapi
C 3	3	794.8	74.3	798	10	BC008261	BC008261 Mus musc
C 4	4	691	64.6	1080	9	AY022624	AY022624 Homo sapi
C 5	5	691	64.6	1210	6	AX464040	AX464040 Sequence
C 6	6	606.6	56.7	814	6	AB040046	AB040046 Sequence
C 7	7	606.6	56.7	814	6	AB015631	AB015631 Sequence
C 8	8	321	30.0	207442	10	AC090489	AB015631 Homo sapi
C 9	9	307	28.7	824	9	BC001027	AC090489 Genomic s
C 10	10	248.6	23.3	101666	2	AC109891	BC001027 Homo sapi
C 11	11	246.8	23.1	649	2	AF186113	AC109891 Rattus no
C 12	12	221	20.7	564	6	AF186113	AF186113 Homo sapi
C 13	13	177.4	16.6	155033	2	AX440375	AX440375 Sequence
C 14	14	177.4	16.6	182350	2	AC025574	AC025574 Homo sapi
C 15	15	177.4	16.6	184762	2	AC073896	AC073896 Homo sapi
C 16	16	126.4	11.8	69753	2	AC103350	AC023500 Homo sapi
C 17	17	88.2	8.3	2378	2	AC103156	AC103156 Rattus no
C 18	18	87.2	8.2	95437	2	AK094445	AK094445 Rattus no
C 19	19	66.8	6.2	147	6	AX096788	AC096788 Rattus no
C 20	20	66.8	6.2	147	6	AX096788	AX096788 Rattus no
C 21	21	56.2	5.3	622	5	AB055671	AX096888 Sequence
C 22	22	52.2	4.9	125020	9	AF429315	AB055671 Datto rer
C 23	23	51.6	4.8	126	6	AX333296	AF429315 Homo sapi
C 24	24	51.6	4.8	126	6	AX333296	AX333296 Sequence
C 25	25	50	4.7	105565	2	AX408687	AX408687 Sequence
C 26	26	49.4	4.6	210631	2	AC129184	AC104103 Mus muscu
C 27	27	49.4	4.6	224840	2	AC027284	AC129184 Mus muscu
C 28	28	49	4.6	2269	10	AF220121	AC027284 Mus muscu
C 29	29	48.6	4.5	168418	9	AC009403	AF220121 Mus muscu
C 30	30	48.4	4.5	218900	2	AC125502	AC009403 Homo sapi
C 31	31	48.2	4.5	2136	2	HSM80446	AC125502 Mus muscu
C 32	32	47.8	4.5	191793	2	AF001897	AL833145 Homo sapi
C 33	33	47.6	4.5	1641	9	BC001744	AF001897 Homo sapi
C 34	34	47.6	4.5	1651	9	BC001744	BC001744 Homo sapi
C 35	35	47.6	4.5	3131	9	HSM803128	BC003108 Homo sapi
C 36	36	47.6	4.5	161566	9	AC008060	BC003108 Homo sapi
C 37	37	47.4	4.4	2326	9	BC011860	AL831915 Homo sapi
C 38	38	47.4	4.4	2435	3	AF007768	AC008060 Homo sapi
C 39	39	47.2	4.4	1583	10	BC034986	BC011860 Homo sapi
C 40	40	47.2	4.4	1688	9	AK026029	AF007768 Choriston
C 41	41	47	4.4	2870	9	HSM800882	BC034986 Mus muscu
C 42	42	47	4.4	167865	2	AC053471	AK026029 Homo sapi
C 43	43	46.8	4.4	1260	6	AX380553	AL110225 Homo sapi
C 44	44	46.8	4.4	1630	6	AX380553	AC053471 Homo sapi
C 45	45	46.8	4.4	169794	2	AC004688	AX380553 Sequence
							BC028028 Homo sapi
							AC004688 Plasmodu

## ALIGNMENTS

RESULT 1	AF186115	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
	AF186115		1017 bp mRNA linear							
				Mus musculus						
				putative secreted protein Zs1g9 (Zs1g9)						
				cda.						
	AF186115									
	AF186115.1									
							GI:6014635			
								Mus musculus.		
								Mus musculus		
								Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
								Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
								1 (bases 1 to 1017)		
								Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and		
								O'Hara,P.		

JOURNAL	Mus musculus putative secreted protein
REFERENCE	2 (bases 1 to 1017)
AUTHORS	Shepard, F., Jelinek, L., Whitmore, T., Blumberg, H., Lehner, J. and O'Hara, P.
TITLE	Direct Submision
JOURNAL	Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA
FEATURES	Location/Qualifiers
source	1..1017
gene	/organism="Mus musculus" /db_xref="dbEST:AA049839" /db_xref="taxon:10090" /clone="IMAGE:476061" 1..1017 /gene="Zs1g9"
CDS	343..891 /gene="Zs1g9" /codon_start=1 /product="putative secreted protein ZS1g9" /protein_id="AA01433.1" /db_xref="GI:6014636" /translation="MKGRGVALLLGLVTGAMARRSODLHCGCRALVDELEWEIAR VDPKRTIOMGSEIRINPDGSOVSVEVPARSAHLTELLEVCQDMKRGQIDPSTHRR KNRYVYVRNSESSELDQGLRIDSIDISGTLKFACSEIVEYEDLEITFEFSREADNYK DKLSKRTDLDHALHRSHEGL"
BASE COUNT	248 a 250 c 315 g 204 t
ORIGIN	
Query Match	95.1%; Score 1017; DB 10; Length 1017;
Best Local Similarity	100.0%; Pred. No. 6e-265;
Matches 1017; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
16	GGGGCTCGCGCCTCGGAGGCGCTCTTAAAGCGCTCGCGCGGCGAGATTGAGG 75
1	GGGGCTCGCGCCTCGGAGGCGCTCTTAAAGCGCTCGCGCGGAGATTGAGG 60
76	GGCGGCTTAGGGCTCAGTTTCGTGGGGGGCTCGCACGGAGCCCTCAGATCTCCGCTTAG 135
61	GGCGGCTTAGGGCTCAGTTTCGTGGGGGGCTCGCACGGAGCCCTCAGATCTCCGCTTAG 120
136	GTGCTTAAAGTCCGGGAAAGCTGGGCGAGCGGTCACTGGCCACCTGAACCTGGCGG 195
121	GTGCTTAAAGTCCGGGAAAGCTGGGCGAGCGGTCACTGGCCACCTGAACCTGGCGG 180
196	GAGCCGGAGCGCTCGGAGAACCGGAGAGCCCGCTTTTCCGAGCCAGCGTAAAGGCT 255
181	GAGCCGGAGCGCTCGGAGAACCGGAGAGCCCGCTTTTCCGAGCCAGCGTAAAGGCT 240
256	TGGGACCCACAGAAACAAAGTAGAGAGTCCGGGCTGCTTCCAGAGCCCTGGGCGAGG 315
241	TGGGACCCACAGAAACAAAGTAGAGAGTCCGGGCTGCTTCCAGAGCCCTGGGCGAGG 300
316	CGGCGGTGGGAGCAGAGTGGAGCGACCCCTGTACACTTAAAGATGAAAGGCTGGGGTTGG 375
301	CGGCGGTGGGAGCAGAGTGGAGCGACCCCTGTACACTTAAAGATGAAAGGCTGGGGTTGG 360
376	CTAGCCCTACTTTTGGGGTCTGCTGTGGAGACTGCGTGGGCTCGAAGAGAGCCAGATGTA 435
361	CTAGCCCTACTTTTGGGGTCTGCTGTGGAGACTGCGTGGGCTCGAAGAGAGCCAGATGTA 420
436	CACGTGGAGCTTGCAGGGGCTCTGGTGGATGAATTAGAGTGGGAAATTTCCCGGTGGAC 495
421	CACGTGGAGCTTGCAGGGGCTCTGGTGGATGAATTAGAGTGGGAAATTTCCCGGTGGAC 480
496	CCCAAGAGACCATTTCAATGGGATCCTTCCGAATCAATCCAGATGGAGCCAGTCAATT 555
481	CCCAAGAGACCATTTCAATGGGATCCTTCCGAATCAATCCAGATGGAGCCAGTCAATT 540
556	GTGAGAGACCTTTATGCCCGCTCAGAGGCCACACGTCACAGAGTTGCTTGAGAGGTGGT 615
541	GTGAGAGACCTTTATGCCCGCTCAGAGGCCACACGTCACAGAGTTGCTTGAGAGGTGGT 600

DB	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
OY	616	GACCGAATGAAGAGATACGGGGAACAGATTGACCTTTACCCACCCGGAAGACTACCTA	675				
DB	601	GACCGAATGAAGAGATACGGGGAACAGATTGACCTTTACCCACCCGGAAGACTACCTA	660				
OY	676	CGCGTGTGAGCCGGGAATGAGAAATCCAGTAACTAGACTTACAGGGCCATCGAATTGAC	735				
DB	661	CGCGTGTGAGCCGGGAATGAGAAATCCAGTAACTAGACTTACAGGGCCATCGAATTGAC	720				
OY	736	TCAGATATACAGCGGACACCTTCAAGTTTCCGTGTGAGAGACATTGTGGAAATACAGAGAT	795				
DB	721	TCAGATATACAGCGGACACCTTCAAGTTTCCGTGTGAGAGACATTGTGGAAATACAGAGAT	780				
OY	796	GAGCTTATCGAATTTCTTCCAGAGAGGCTCAACAAGTTAAAGACAACTTTGCAAGTAA	855				
DB	781	GAGCTTATCGAATTTCTTCCAGAGAGGCTCAACAAGTTAAAGACAACTTTGCAAGTAA	840				
OY	856	CGGACAGATCTATGTGACCATGCCCTGCAGAGATCTCAGATGAGCTATGATCACTGGA	915				
DB	841	CGGACAGATCTATGTGACCATGCCCTGCAGAGATCTCAGATGAGCTATGATCACTGGA	900				
OY	916	GCAAGCACCCTTACCAACCAAGTGTGAAACACCCCGAGAGGGGAAGATGCGACATTGC	975				
DB	901	GCAAGCACCCTTACCAACCAAGTGTGAAACACCCCGAGAGGGGAAGATGCGACATTGC	960				
OY	976	CTTTATATTTACGTTTTATGAAATGAACAGAAAAAAGCTTTGAAACCGAAAGTA	1032				
DB	961	CTTTATATTTACGTTTTATGAAATGAACAGAAAAAAGCTTTGAAACCGAAAGTA	1017				
RESULT 2	AC012013	158198 bp	DNA	linear	HTG 17-JUL-2001		
AC012013	LOCUS	Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT					
DEFINITION	SEQUENCE, 5 unordered pieces.						
AC012013	AC012013						
VERSION	AC012013.18	GI:14717292					
KEYWORDS	HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.						
SOURCE	Homo sapiens.						
ORGANISM	Homo sapiens.						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
AUTHORS	Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C., Alshrooke,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbata,J., Benton,J., Bilmage,K., Blankenburg,K., Bonnin,D., Boucek,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Butcher,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Demn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnheart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodson,A., Hogues,M., Holloway,C., Hollins,B., Homel,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,Z., Kureshi,A., Landry,N., Leel,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Lolesleged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Mbhehwarl,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mel,G., Metker,M., Miner,G., Miner,Z., Mitchell,T., Monabat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newson,J., Newson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenko,S., Ogunu,M., Okunoye,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Peters,L., Peters,L., Plickens,R., Plims,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Saveri,G., Scherer,S., Scott,G., Shen,H., Shooshbari,N.,						



[illegible]

Accession	Gene	Species	Chromosome	Start (kb)	End (kb)	Orientation	Length (bp)	GC Content (%)	GC Content (3bp window)	GC Content (6bp window)	GC Content (12bp window)	GC Content (24bp window)	GC Content (48bp window)	GC Content (96bp window)	GC Content (192bp window)	GC Content (384bp window)	GC Content (768bp window)	GC Content (1536bp window)	GC Content (3072bp window)	GC Content (6144bp window)	GC Content (12288bp window)	GC Content (24576bp window)	GC Content (49152bp window)	GC Content (98304bp window)	GC Content (196608bp window)	GC Content (393216bp window)	GC Content (786432bp window)	GC Content (1572864bp window)	GC Content (3145728bp window)	GC Content (6291456bp window)	GC Content (12582912bp window)	GC Content (25165824bp window)	GC Content (50331648bp window)	GC Content (100663296bp window)	GC Content (201326592bp window)	GC Content (402653184bp window)	GC Content (805306368bp window)	GC Content (1610612736bp window)	GC Content (3221225472bp window)	GC Content (6442450944bp window)	GC Content (12884901888bp window)	GC Content (25769803776bp window)	GC Content (51539607552bp window)	GC Content (103079215104bp window)	GC Content (206158430208bp window)	GC Content (412316860416bp window)	GC Content (824633720832bp window)	GC Content (1649267441664bp window)	GC Content (3298534883328bp window)	GC Content (6597069766656bp window)	GC Content (13194139533312bp window)	GC Content (26388279066624bp window)	GC Content (52776558133248bp window)	GC Content (105553116266496bp window)	GC Content (211106232532992bp window)	GC Content (422212465065984bp window)	GC Content (844424930131968bp window)	GC Content (1688849860263936bp window)	GC Content (3377699720527872bp window)	GC Content (6755399441055744bp window)	GC Content (13510798882111488bp window)	GC Content (27021597764222976bp window)	GC Content (54043195528445952bp window)	GC Content (108086391056891840bp window)	GC Content (216172782113783680bp window)	GC Content (432345564227567360bp window)	GC Content (864691128455134720bp window)	GC Content (1729382256910269440bp window)	GC Content (3458764513820538880bp window)	GC Content (6917529027641077760bp window)	GC Content (13835058055282155520bp window)	GC Content (27670116110564311040bp window)	GC Content (55340232221128622080bp window)	GC Content (110680464442257244160bp window)	GC Content (221360928884514488320bp window)	GC Content (442721857769028976640bp window)	GC Content (885443715538057953280bp window)	GC Content (1770887431076115906560bp window)	GC Content (3541774862152231813120bp window)	GC Content (7083549724304463626240bp window)	GC Content (14167099448608927252480bp window)	GC Content (28334198897217854504960bp window)	GC Content (56668397794435709009920bp window)	GC Content (113336795588871418019840bp window)	GC Content (226673591177742836039680bp window)	GC Content (453347182355485672079360bp window)	GC Content (906694364710971344158720bp window)	GC Content (1813388729421942688317440bp window)	GC Content (3626777458843885376634880bp window)	GC Content (7253554917687770753269760bp window)	GC Content (14507109835375541506539520bp window)	GC Content (29014219670751083013079040bp window)	GC Content (58028439341502166026158080bp window)	GC Content (116056878683004332052316160bp window)	GC Content (232113757366008664104632320bp window)	GC Content (464227514732017328209264640bp window)	GC Content (9284550294640346564185281280bp window)	GC Content (18569100589280693123770562560bp window)	GC Content (37138201178561386247541125120bp window)	GC Content (74276402357122772495108250240bp window)	GC Content (14855280471424554990221650480bp window)	GC Content (29710560942849109980443300960bp window)	GC Content (59421121885698219960886601920bp window)	GC Content (118842243771396439921773233840bp window)	GC Content (237684487542792879843546467680bp window)	GC Content (475368975085585759687092935360bp window)	GC Content (950737950171171519374185870720bp window)	GC Content (1901475900342343038748371741440bp window)	GC Content (3802951800684686077496743482880bp window)	GC Content (7605903601369372154993486965760bp window)	GC Content (15211807202738744309986939331520bp window)	GC Content (30423614405477488619973878663040bp window)	GC Content (60847228810954977239947757326080bp window)	GC Content (121694457621909954479895514652160bp window)	GC Content (243388915243819908959791029304320bp window)	GC Content (486777830487639817919582058608640bp window)	GC Content (973555660975279635839164117217280bp window)	GC Content (1947111321950559271678328234434560bp window)	GC Content (3894222643901118543356656468869120bp window)	GC Content (7788445287802237086713312937738240bp window)	GC Content (15576890575604474173426625875476480bp window)	GC Content (31153781151208948346853251750952960bp window)	GC Content (62307562302417896693706503501905920bp window)	GC Content (124615124604835793387413007003811840bp window)	GC Content (249230249209671586774826014007623680bp window)	GC Content (498460498419343173549652028015247360bp window)	GC Content (996920996838686347099304056030494720bp window)	GC Content (19938419936773726941986081120118948480bp window)	GC Content (3987683987354745388397216224038989760bp window)	GC Content (7975367974709490776794432448077979520bp window)	GC Content (15950735949418981553588864896155599040bp window)	GC Content (31901471898837963107177729792319198080bp window)	GC Content (63802943797675926214355459584638396160bp window)	GC Content (1276058875953518524287109191172767932320bp window)	GC Content (255211775190703704857421882344554864640bp window)	GC Content (510423550381407409714843644689109729280bp window)	GC Content (102084710076281481942
-----------	------	---------	------------	------------	----------	-------------	-------------	----------------	-------------------------	-------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------------------------	---------------------------	---------------------------	----------------------------	----------------------------	----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	------------------------------	------------------------------	------------------------------	-------------------------------	-------------------------------	-------------------------------	--------------------------------	--------------------------------	--------------------------------	---------------------------------	---------------------------------	---------------------------------	---------------------------------	----------------------------------	----------------------------------	----------------------------------	-----------------------------------	-----------------------------------	-----------------------------------	------------------------------------	------------------------------------	------------------------------------	------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	--------------------------------------	--------------------------------------	--------------------------------------	---------------------------------------	---------------------------------------	---------------------------------------	---------------------------------------	--	--	--	---	---	---	--	--	--	--	---	---	---	--	--	--	---	---	---	---	--	--	--	---	---	---	--	--	--	--	---	---	---	--	--	--	---	---	---	--	---	---	---	---	---	---	--	--	--	--	---	---	---	--	--	--	---	---	---	---	--	--	--	---	---	---	--	--	--	--	--	---	---	--	--	--	--	---	---	-----------------------------------

COMMENT Contact: MGC help desk  
Email: cga@bbs-triemail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: amg@bcm.tmc.edu  
Guanarane, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>  
Series: IRAK Plate: 5 Row: b Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA g1: 9903606.  
Location/Qualifiers  
1. 798  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="129, C57BL/6J, FVB/N"  
/clone="MGC:6853 IMAGE:2650612"  
/tissue="MGC:6853 IMAGE:2650612"  
months old, gross tissue.  
/tissue\_type="mammary tumor. Brcal-/fl: MMTV-Cre model. 10  
months old, gross tissue."  
/clone\_11b="NCI\_CGAP\_Mam3"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
115. 663  
/codon\_start=1  
/product="transmembrane protein 4"  
/protein\_id="AA08261.1"  
/db\_xref="GI:14198401"  
/db\_xref="LocusID:56530"  
/translation="MKMGWIALLLGALITGNARRSODLHCACRALVDELEMIAR  
VDPKKTQMGSRINPDGSGSVSEVPPARSBAHUTELLEECDDMKKEGQIDPSTR  
KNYRVVGRNGESSELDLQIRIDSDISGLTKFACESIVEEDELLEFFSRADNVK  
DKLCSKRTDLCDAHLHRSDEL"

BASE COUNT 227 a 182 c 228 g 161 t  
ORIGIN

Query Match 74.3%; Score 794.8; DB 10; Length 798;  
Best Local Similarity 99.7%; Pred. No. 1.2e-204;  
Matches 796; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 244 AGCTCTAGAGGTGGGACCCACAGAAACAAAGTGAAGTCCGGCTGCTTCACAGCCT 303  
1 ACCTCTAGAGGTGGGACCCACAGAAACAAAGTGAAGTCCGGCTGCTTCACAGCCT 60  
QY 304 GGGCCACGGCGGCGCCGCTGGGAGCAGAGGTGAGCAGCCTGTTACACTAAAGTGA 363  
61 GGGCCACGGCGGCGCCGCTGGGAGCAGAGGTGAGCAGCCTGTTACACTAAAGTGA 120  
Db 364 GGGTGGGGTGGCTAGCCCTACTTTTGGGGCTCTGCTGGGAACCTGCTGGGCTGGAAG 423  
121 GGGTGGGGTGGCTAGCCCTACTTTTGGGGCTCTGCTGGGAACCTGCTGGGCTGGAAG 180  
QY 424 AGCCAAAGATCTTACATCTTGAGAGCTTGAAGGGCTCTGATGAATTAAGTGGGAAT 483  
181 AGCCAAAGATCTTACATCTTGAGAGCTTGAAGGGCTCTGATGAATTAAGTGGGAAT 240  
Db 484 GCCCGCGTGGACCCCAAGAAAGCATTCAGATGGATCCTTCGATCAATCAGATGGC 543  
241 GCCCGCGTGGACCCCAAGAAAGCATTCAGATGGATCCTTCGATCAATCAGATGGC 300  
QY 544 AGCCAGTAGTTTGGAGAGTATTATGCGGCTCAGAGGCCACCTCAGACAGATTGCT 603  
301 AGCCAGTAGTTTGGAGAGTATTATGCGGCTCAGAGGCCACCTCAGACAGATTGCT 360  
Db 604 GAGGAGGTGTGACCCGAATGAAGAGTACGGGGAACAGATTGACCTTTCACCAACGC 663  
|||||

Db 361 GAGAGAGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCTTTCACCAACGC 420  
QY 664 AAGAACTACGTACGCGTCTGTGAGCCGGAATGGAATCCAGTGAATCAGTACAGGCG 723  
|||||  
Db 421 AAGAACTACGTACGCGTCTGTGAGCCGGAATGGAATCCAGTGAATCAGTACAGGCG 480  
QY 724 AACCGAATGACTCAGATATACAGCGGACCCCTCAAGTTTGCCTGTGAGAGCATTTGGA 783  
784 AACCGAATGACTCAGATATACAGCGGACCCCTCAAGTTTGCCTGTGAGAGCATTTGGA 843  
481 AACCGAATGACTCAGATATACAGCGGACCCCTCAAGTTTGCCTGTGAGAGCATTTGGA 540  
Db 541 GAATACGAGATGAGCTTATCGAATTTCTTCACAGAGAGCTGACACCTTAAACACAA 600  
QY 844 CTTTGAGTAGCAGCAGACAGATATATGACATGCGCTCAGACAGATTCACAGTACGTA 903  
601 CTTTGAGTAGCAGCAGACAGATATATGACATGCGCTCAGACAGATTCACAGTACGTA 660  
QY 904 TGAATCTACTGAGCAGACAGCCTTACACCAACCTGATGAGACACCCCGAGAGGGA 963  
661 TGAATCTACTGAGCAGACAGCCTTACACCAACCTGATGAGACACCCCGAGAGGGA 720  
QY 964 TGGCAGCATTCGCTTTATATATACGTTTATGAAATGAACTGAAATACCTTGA 1023  
721 TGGCAGCATTCGCTTTATATATACGTTTATGAAATGAACTGAAATACCTTGA 780  
Db 1024 CCGAAAGTAAAAA 1041  
781 CCGAAAGTAAAAA 798  
Db 781 CCGAAAGTAAAAA 798

RESULT 4  
AY032624 1080 bp mRNA linear PRI 18-APR-2002  
LOCUS  
DEFINITION Homo sapiens saposin-like protein mRNA, complete cds.  
ACCESSION AY032624  
VERSION  
KEYWORDS  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1080)  
Bornhauser, B.C., Olsson, P.-A. and Lindholm, D.  
NSAP is a novel saposin-like protein that interacts with MIR and  
stimulates neurite outgrowth  
unpublished  
REFERENCE 2 (bases 1 to 1080)  
Olsson, P.-A. and Lindholm, D.  
Direct Submission  
Submitted (13-APR-2001) Neuroscience, Uppsala University,  
Husargatan 3, Uppsala 75123, Sweden  
LOCATION/Qualifiers  
1. 1080  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q15"  
419. 967  
/note="NSAP: encodes a type 2 membrane signal and a  
putative ER retention motif"  
/product="saposin-like protein"  
/protein\_id="AAK38148.1"  
/db\_xref="GI:20196199"  
/translation="MKMGWIALLLGALITGNARRSODLHCACRALVDELEMIAR  
VDPKKTQMGSRINPDGSGSVSEVPPARSBAHUTELLEECDDMKKEGQIDPSTR  
KNYRVVGRNGESSELDLQIRIDSDISGLTKFACESIVEEDELLEFFSRADNVK  
DKLCSKRTDLCDAHLHRSDEL"

BASE COUNT 238 a 269 c 343 g 230 t  
ORIGIN

Query Match 64.6%; Score 691; DB 9; Length 1080;



```

OY 669 CTACGTACCGCTGTCGAGCCGGAATGAGAAATCCAGTGAACCTTACAGGCGCATCGG 728
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 854 CTACGTACCGCTGTCGAGCCGGAATGAGAAATCCAGTGAACCTTACAGGCGCATCGG 913
OY 729 AATTGACCTAGATATGAGCGGACCCCTCAAGTTTGGTGTGAGAGATGTGGAAAGAA 788
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 914 AATTGACCTAGATATGAGCGGACCCCTCAAGTTTGGTGTGAGAGATGTGGAAAGAA 973
OY 789 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGCGCTGCAACGTTAAAGACAACTTTG 848
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 974 CGAGATGAGCTTATGCAATTTCTTCCAGAGAGCGCTGCAACGTTAAAGACAACTTTG 1033
OY 849 CAGTGAAGCGGACAGATCTATGTGACCATGCTCCGACAGATCTCAGATGAGCTATGAAT 908
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1034 CAGTGAAGCGGACAGATCTATGTGACCATGCTCCGACAGATCTCAGATGAGCTATGAAC 1093
OY 909 CACTGAGACGACGCTACACCAACGTCATGGAACACCCCGAGAGGGGAAAGATGGA 968
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1094 CACTGAGACGACCCAC-----ACTGCTTGATGATGATCACCACGAGAGGGGAAATGGA 1148
OY 969 GCATGCGCTTT--TATATTACGTTTTTATGAAATGAACTGAAAAA 1013
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1149 GCAATGCGCTTTTATATATATATGTTTTTACTGAAATTAAGTGAAGAAA 1195

RESULT 6
AA440456 814 bp DNA linear PAT 28-JUN-2002
LOCUS Sequence 309 from Patent WO0190154.
DEFINITION AX440456
ACCESSION AX440456
VERSION AX440456.1 GI:2165266
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Xu,J., Mitcham,J.L., Harlocker,S.L., Dillon,D.C., Secrett,H.,
Lodes,M.J., Algate,P.A., Fling,S.P., Mannion,J., Benson,D.R. and
Carter,D.
TITLE
Compositions and methods for the therapy and diagnosis of ovarian
cancer
JOURNAL
Patent: WO 0190154-A 309 29-NOV-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
source
1..814
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 210 a 194 c 241 g 169 t
ORIGIN
Query Match 56.7%; Score 606.6; DB 6; Length 814;
Best Local Similarity 87.3%; Pred. No. 1.5e-153;
Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;
OY 213 AGAAGCCGGGACAGCCCGTTTCCAGCCAGCTGCTAGGGTGGAGCCACAGAAAC 272
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1 AGAATCCCGGACAGCCCTGCTCCCTGACAGCGTG-TAGTTTCGGAGACCACTGGGCG 59
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
OY 273 AAAGTGAAGTCCGGCTCTCCAGAGCGTGGGACGCGCGCGCGCTGGAGCAGAG 332
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 60 AAAGTGAAGTCCAGCGGCTCTCCAGCGCTTGGGACGCGCGCGCTGGAGCAGAG 119
OY 333 GTGAGCAGACCTGTTACATAAAGATGAAGGCTGGGGTGGTGAAGCTTCTTTGGG 392
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 120 GTGAGCAGACCTGTTACATAAAGATGAAGGCTGGGGTGGTGAAGCTTCTTTGGG 179
OY 333 GGTCTGCTGGGAATGCTGCTGGGCTGGAAGGACCAAGATCTACCTGTGAGCTTGCAG 452
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 180 GGCCTCTGTGGAAACCGCTGGGCTGCGAGGAGCAGATCTCCACTGTGAGCATGCGAG 239
OY 453 GGCCTCTGTGGATGAATTAAGTGGAAATTCGCCGCGTGGAGCCCAAGAACCATTTCA 512

```

```

DB 240 GGCCTCTGTGGATGAATTAAGTGGAAATTCGCCGCGTGGAGCCCAAGAACCATTTCA 299
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
OY 513 GATGGATCTCTCGAATCAATCCAGATGGAGCCGATGAGTTGTGGAGTTACCTTATGC 572
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 300 GATGGATCTCTCGAATCAATCCAGATGGAGCCGATGAGTTGTGGAGTTACCTTATGC 359
OY 573 CCGCTCAGAGGCCACCTTCACAGAGTTGCTTGAGAGGTGTGTGACCGAATGAAGAGTA 632
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 360 CCGCTCAGAGGCCACCTTCACAGAGTTGCTTGAGAGGTGTGTGACCGAATGAAGAGTA 419
OY 633 CCGGGAACAGATTGATCTTCCACCATGCAAGAACTGATGATGATGATGATGATGATG 479
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 420 TGGGGAACAGATTGATCTTCCACCATGCAAGAACTGATGATGATGATGATGATGATG 479
OY 693 TGGAGAAATCCAGTAATCTTACAGGCAATCCGAATGATGATGATGATGATGATGATG 752
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 480 TGGAGAAATCCAGTAATCTTACAGGCAATCCGAATGATGATGATGATGATGATGATG 539
OY 753 CCTCAAGTTTGGCTGTGAGAGCATTTGTGAGAGAAATGAGAGATGATGATGATGATG 812
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 540 CCTCAAGTTTGGCTGTGAGAGCATTTGTGAGAGAAATGAGAGATGATGATGATGATG 599
OY 813 CTCCAGAGAGGCTGACACATTAAGACAACTTTGCACTAAGCGGACAGATCTATGTA 872
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 600 TTCCCGAGAGGCTGACAAATGTAAGACAACTTTGCACTAAGCGGACAGATCTTTGTA 659
OY 873 CCATGCGCTGACAGATCTCAGATGATGATGATGATGATGATGATGATGATGATGATG 932
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 660 CCATGCGCTGACAGATCTCAGATGATGATGATGATGATGATGATGATGATGATGATG 714
OY 933 AACGTGATGGAACACCCCGAGAGGGAAGATGCGACATTTGCTTT--TATATTACGTT 990
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 715 GCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774
OY 991 TTTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 774
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 775 TTTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 797

RESULT 7
AB015631 814 bp mRNA linear PRI 26-JUL-2001
LOCUS Homo sapiens mRNA for type II membrane protein, complete cds,
DEFINITION clone:HP10390.
ACCESSION AB015631
VERSION AB015631.1 GI:4586839
KEYWORDS type II membrane protein.
SOURCE Homo sapiens gastric adenocarcinoma CDNA to mRNA,
clone:lib:pkal-meta-1 clone:HP10390.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Yokoyama-Kobayashi,M., Yamaguchi,T., Sekine,S. and Kato,S.
TITLE
Selection of cDNAs encoding putative type II membrane proteins on
the cell surface from a human full-length cDNA bank
JOURNAL
Gene 228 (1-2), 161-167 (1999)
MEDLINE 99173880
REFERENCE
2 (bases 1 to 814)
AUTHORS Kato,S.
TITLE
Direct Submission
JOURNAL
Submitted (16-JUN-1998) Selshi Kato, Research Institute of National
Rehabilitation Center for the Disabled, Department of
Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama
359-8555, Japan (E-mail:selshi@rephab.go.jp
Tel:042-995-3100(ex.2568), Fax:042-995-3132)
FEATURES
Location/Qualifiers
source
1..814
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HP10390"
/tissue_type="gastric adenocarcinoma"

```

CDS /clone\_11b="pkal-meta-1"

145..693

/codon\_start=1  
/product="type II membrane protein"

/protein\_id="BAA76498.1"

/db\_xref="GI:4586840"

/translation="KKGWMLLLGLLGTAMARRSODLHGCACRAVLDELEWIAQ  
VDPKTIQMSFRINPDSQSVPEVPAKSEHLEICDPAKREKGEQIDPSTHR  
KNYRVVNGESSELDLQIRIDISDLSGLTFACESEIVEEEDLIEFFSEADNVK  
DKLSKRDLCDHHAHSHDEL"

BASE COUNT 210 a 194 c 241 g 169 t

# Query Match

Best Local Similarity 87.3%; Score 606.6; DB 9; Length 814;  
Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;

OY 213 AGAAGCCGGGAGACCCCTTTTCCACAGCGCTAGGCTGGGAGCCACAGAAAC 272  
DB 1 AGAATCCCGAGAGCCCTCTCCCTGACAGCGTG-1AGTTCCGGAGCCACTGGGGCC 59  
OY 273 AAGTGAAGTCCGGCTGTTTCCAGAGCTTGGCCAGCGCGCGCTGGAGACAG 332  
DB 60 AAGTGAAGTCCGGCTGTTTCCAGAGCTTGGCCAGCGCGCGCGCTGGAGACAG 119  
OY 333 GTGAGAGCAGCCTGTACCTAAGATGAAGGCTGGGCTGGCTAGCCCTTACTTTGGG 392  
DB 120 GTGAGAGCAGCCTGTACCTAAGATGAAGGCTGGGCTGGCTAGCCCTTACTTTGGG 179  
OY 393 GGTCTGCTGGGAGTCCCTGGCTCGAAGAGCCAGATCTACCTGTGAGAGCTTGGAG 452  
DB 180 GGCCTGCTGGGAGTCCCTGGCTCGAAGAGCCAGATCTACCTGTGAGAGCTTGGAG 239  
OY 453 GGCCTGCTGGGAGTCCCTGGCTCGAAGAGCCAGATCTACCTGTGAGAGCTTGGAG 512  
DB 240 GGCCTGCTGGGAGTCCCTGGCTCGAAGAGCCAGATCTACCTGTGAGAGCTTGGAG 299  
OY 513 GATGGAGATCTCCGGAATCAATCCAGATGCGAGCAGTCACTGTGGAGGTAAGCTTATGC 572  
DB 300 GATGGAGATCTCCGGAATCAATCCAGATGCGAGCAGTCACTGTGGAGGTAAGCTTATGC 359  
OY 573 CCGCTCAGAGCCCACTCAGAGAGTGTGAGAGAGTGTGAGAGAGTGAAGAGTGA 632  
DB 360 CCGCTCAGAGCCCACTCAGAGAGTGTGAGAGAGTGTGAGAGAGTGAAGAGTGA 419  
OY 633 CGGGAACAGATTGACCTTCTACCCAGCCGGAAGTACAGTCCGTCTGAGCCGGA 692  
DB 420 TGGGGAACAGATTGATCTTCCACCCATCCGAGAACTACGTAGTGGGCGGAA 479  
OY 693 TGGGAATCCAGTGAATCAATCCAGAGTCCGGAATGAGTCAAGATATCAAGCGGAC 752  
DB 480 TGGGAATCCAGTGAATCAATCCAGAGTCCGGAATGAGTCAAGATATCAAGCGGAC 539  
OY 753 CCTCAAGTGTGCGTGTGAGAGCAATGGGAAGTGAAGAGTGAAGTGAAGTGA 812  
DB 540 CCTCAAGTGTGCGTGTGAGAGCAATGGGAAGTGAAGAGTGAAGTGAAGTGA 599  
OY 813 CTCAAGAGAGGTGAGCAACGTTAAAGCAAACTTTCAGTGAAGAGAGTGAAGTGA 872  
DB 600 TTTCCAGAGAGGTGAGCAACGTTAAAGCAAACTTTCAGTGAAGAGAGTGAAGTGA 659  
OY 873 CCAATGCGCTGACAGATTCACAGATGATGATGATGATGATGATGATGATGATGATGAT 932  
DB 660 CCAATGCGCTGACAGATTCACAGATGATGATGATGATGATGATGATGATGATGATGAT 714  
OY 933 AAGCTGATGAACACCCAGAGAGGAGATGAGCAGATGAGCTTATATATATATATATAT 990  
DB 715 GAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774  
OY 991 TTTATGGAATGAATGAAAAA 1013  
DB 775 TTTACTGAATTAATGAAAAA 797

## RESULT 8

AC090489

LOCUS

DEFINITION

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489 207424 bp DNA linear ROD 26-JAN-2002

Genomic sequence for Mus musculus, clone RP23-104010, complete

sequence.

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

AC090489

|||||  
DB 72364 TAGCCGAGACGCTCTGAGAAAGCCGAGACGCCCTTTTCCACAGCTGCTAGGGT 72423  
OY 256 TGGAGCCACACAGAAAACAAAGTGAAGTCCGCTGCTTTCAGAGCCTGGCCACGCGG 315  
DB 72424 TGGAGCCACACAGAAAACAAAGTGAAGTCCGCTGCTTTCAGAGCCTGGCCACGCGG 72483  
OY 316 CGGCGCTGGAGACAGAGTGG 336  
DB 72484 CGGCGCTGGAGACAGAGTGG 72504  
RESULT 9  
BC001027  
LOCUS BC001027 824 bp mRNA linear PRI 12-JUL-2001  
DEFINITION Homo sapiens, transmembrane protein 4, clone MGC:1545  
IMAGE:344788, mRNA, complete cds.  
ACCESSION BC001027  
VERSION BC001027.1 GI:12654402  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 824)  
Strausberg, R.  
Direct Submission  
Submitted (17-NOV-2000) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: http://mgc.ncl.nih.gov/  
Contact: MGC help desk  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: http://www.nisc.nih.gov/  
Contact: nisc.mgc@nih.gov  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J.,  
Tlonsong, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L.-H. and Green, E.D.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: http://image.liml.gov  
Series: IMAGE Plate: 5 Row: P Column: 13  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 6014631.  
FEATURES  
SOURCE  
1. 824  
/organism="Homo sapiens"  
/db\_xref="locusid:10330"  
/db\_xref="taxon:9606"  
/clone="MGC:1545 IMAGE:344788"  
/tissue\_type="Eye, retinoblastoma"  
/clone\_id="NH\_MGC\_16"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
250..504  
/codon\_start=1  
/product="transmembrane protein 4"  
/protein\_id="AA01027.1"  
/db\_xref="GI:12654403"  
/translacion="MKGMGLALIGALGTAMARRSODLHCACRALVDELEMEIAQ  
VDPKKTIONMSFRNPDSOSVETVTVPNNKVAHSGFG"  
BASE COUNT 221 a 194 c 233 g 176 t

ORIGIN  
Query Match 28.7%; Score 307; DB 9; Length 824;  
Best Local Similarity 82.9%; Pred. No. 3.e-72;  
Matches 374; Conservative 0; Mismatches 75; Indels 2; Gaps 2;  
OY 123 GATCTCCGCTTAGGTGCTTACGTTAGTGGGGAAGTGGGCCAGCGGTC-ACGGCCAC 181  
DB 15 GAGTCCCGCTTAGGTGCTTACGTTAGTGGGGAAGTGGGCCAGCGGTCAGCGGACAC 74  
OY 182 CCGTAACCTGGGGGAGCGCGGCTTGGGAAGCGGGGACAGCGCGCTTTTCCCGAG 241  
DB 75 CCGGTGTGGCAGAGCGCGGAGAGCGCTTGGGAATCCGGAGCGCTGCTCCGAG 134  
OY 242 CCAGCTGTAGGGTGGGAGCCACAGAAAACAAAGTGAAGTCCGCTCTTCCAGAC 301  
DB 135 CCAGGTG-TAGTTTGGGAGCCACAGTGGGCGAAAGTGAAGTCCAGGGTCTTCCAGCGC 193  
OY 302 CTGGGCCACGCGCGCGCGCGCTGGGAGCAGAGTGGAGCCACCTGTTACATTAAGATGA 361  
DB 194 TTGGGCCACGCGCGCGCGCGCTGGGAGCAGAGTGGAGCCACCTATTAAGATGA 253  
OY 362 AAGGCTGGGGGTGGCTAGCGCTACTTTGGGGGTCTGCTGGGAAGTGGCTGGCGA 421  
DB 254 AAGGCTGGGGGTGGCTAGCGCTACTTTGGGGGTCTGCTGGGAAGTGGCTGGCGA 313  
OY 422 GGAGCCAGATCTACACTGTGGAGCTTGCAGGGCTCTGTGATGAATAGAGTGGGAA 481  
DB 314 GGAGCCAGATCTACACTGTGGAGCTTGCAGGGCTCTGTGATGAATAGAGTGGGAA 373  
OY 482 TTGGCCGCTGAGACCCCAAGAACCATTCAGATGGATGCTTCCGAATCAATCCAGATG 541  
DB 374 TTGGCCGCTGAGACCCCAAGAACCATTCAGATGGATGCTTCCGAATCAATCCAGATG 433  
OY 542 GCAGCAGTCACTGTGGAGTACCTTATGAC 572  
DB 434 GCAGCAGTCACTGTGGAGTACCTTATGAC 464  
RESULT 10  
AC109891 10166 bp DNA linear HTG 13-JUL-2002  
LOCUS AC109891  
DEFINITION Rattus norvegicus clone CH230-319E6, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 47 unordered pieces.  
AC109891  
VERSION AC109891.3 GI:21738196  
KEYWORDS HTG; HTGS; PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 10166)  
Munry, D.M., Adams, C., Adio-Oduola, B., Ali-oshan, F.R., Allen, C.,  
Alibrooks, S.L., Amaralunga, H.C., Are, J.R., Ayala, M., Banks, T.,  
Barbarta, J., Benton, J., Blmage, K., Blanchenburg, K., Bonnin, D.,  
Bouck, J., Bowler, S., Brileva, M., Brown, M., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carroll, T.F., Carter, M., Cavazos, R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dalhorne, S.R., David, R.,  
DeVella, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earnhardt, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, D., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorell, J.D., Guvvara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Honsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivel, S., Joudah, S.,  
Karlisson, E., Kelly, S., Khan, U., King, L., Korvah, D., Kovar, C.,



Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louieged, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, J., Mohabbat, K., Morgan, K., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newson, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwunonu, G.,  
Oragunye, N., Oriedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivera, M., Rojas, A., Rojudo, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scheer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansley, J., Taylor, C., Taylor, T., Teitrod, B., Thomas, N., Thomas, S.,  
Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 101666)  
Morley, K.C.  
Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 101666)  
Morley, K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced g1:18860222.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GSKY  
Center clone name: CH230-319P6  
----- Summary Statistics  
Sequencing vector: Plasmid  
Assembly: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 60229 bases at least Q40  
Consensus quality: 63959 bases at least Q30  
Consensus quality: 66730 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 47 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1423: contig of 1423 bp in length  
\* 1424 1523: gap of unknown length  
\* 1524 2972: contig of 1449 bp in length  
\* 2973 3072: gap of unknown length  
\* 3073 4717: contig of 1645 bp in length  
\* 4718 4817: gap of unknown length  
\* 4818 6632: contig of 1815 bp in length  
\* 6633 6732: gap of unknown length  
\* 6733 8040: contig of 1308 bp in length  
\* 8041 8140: gap of unknown length  
\* 8141 10043: contig of 1903 bp in length  
\* 10044 10143: gap of unknown length

10144 11237: contig of 1094 bp in length  
\* 11238 11337: gap of unknown length  
\* 11338 12362: contig of 1025 bp in length  
\* 12363 12462: gap of unknown length  
\* 12463 14100: contig of 1538 bp in length  
\* 14101 14201: gap of unknown length  
\* 14202 15840: contig of 1640 bp in length  
\* 15841 15940: gap of unknown length  
\* 15941 17053: contig of 1113 bp in length  
\* 17054 17154: gap of unknown length  
\* 17155 18836: contig of 1683 bp in length  
\* 18837 18937: gap of unknown length  
\* 18938 20137: contig of 1201 bp in length  
\* 20138 20237: gap of unknown length  
\* 20238 21736: contig of 1499 bp in length  
\* 21737 21837: gap of unknown length  
\* 21838 23620: contig of 1783 bp in length  
\* 23621 25907: gap of unknown length  
\* 25908 26007: contig of 2188 bp in length  
\* 26008 27407: gap of unknown length  
\* 27408 27506: contig of 1399 bp in length  
\* 27507 28837: gap of unknown length  
\* 28838 28937: contig of 1331 bp in length  
\* 28938 30230: gap of unknown length  
\* 30231 30330: contig of 1292 bp in length  
\* 30331 32013: gap of unknown length  
\* 32014 32113: contig of 1683 bp in length  
\* 32114 33198: gap of unknown length  
\* 33199 33298: contig of 1086 bp in length  
\* 33299 35519: gap of unknown length  
\* 35520 35619: contig of 2221 bp in length  
\* 35620 37678: gap of unknown length  
\* 37679 37777: contig of 2038 bp in length  
\* 37778 38936: gap of unknown length  
\* 38937 39036: contig of 1159 bp in length  
\* 39037 40273: gap of unknown length  
\* 40274 40373: contig of 1237 bp in length  
\* 40374 41906: gap of unknown length  
\* 41907 42006: contig of 1533 bp in length  
\* 42007 44506: gap of unknown length  
\* 44507 44606: contig of 2500 bp in length  
\* 44607 46803: gap of unknown length  
\* 46804 46903: contig of 2197 bp in length  
\* 46904 49326: gap of unknown length  
\* 49327 49427: contig of 2323 bp in length  
\* 49428 51433: gap of unknown length  
\* 51434 51533: contig of 2107 bp in length  
\* 51534 53763: gap of unknown length  
\* 53764 53863: contig of 2230 bp in length  
\* 53864 55613: gap of unknown length  
\* 55614 56713: contig of 2750 bp in length  
\* 56714 58773: gap of unknown length  
\* 58774 58873: contig of 2060 bp in length  
\* 58874 61527: gap of unknown length  
\* 61528 61627: contig of 2654 bp in length  
\* 61629 63192: gap of unknown length  
\* 63193 63292: contig of 1565 bp in length  
\* 63293 65362: gap of unknown length  
\* 65363 65462: contig of 2070 bp in length  
\* 65463 67438: gap of unknown length  
\* 67439 67538: contig of 1976 bp in length  
\* 67539 70505: gap of unknown length  
\* 70506 70605: contig of 2967 bp in length  
\* 70606 74524: gap of unknown length  
\* 74525 74624: contig of 3919 bp in length  
\* 74625 76737: gap of unknown length  
\* 76738 76837: contig of 2113 bp in length  
\* 76838 80133: gap of unknown length  
\* 80134 80233: contig of 3296 bp in length  
\* 80234 83122: gap of unknown length  
\* 83123 83222: contig of 2889 bp in length  
\* 83223 86731: gap of unknown length  
\* 86732 86733: contig of 3509 bp in length





Matches	298	Conservative	0	Mismatches	50	Indels	10	Gaps	4
OY	211	GGAGAACCCGGGACAGCCCGCTTTTCCAGCAGCTGCTAGGGTGGGAGCCACAGAAA	270						
Db	216	GGAGAAATCCCGACAGCCCTGCTCCCTCAGCCAGAGTGTGTTTGGGAGCCACTGGG	274						
OY	271	ACAAGTAGAGTCCGGCTGCTTCCAGAGCCTGGGCGGCGGCGGCTGGAGACAG	330						
Db	275	CCAAAGTAGAGTCCAGGCTGCTTCCAGAGCCTGGGCGGCGGCGGCTGGAGACAG	334						
OY	331	AGGTGAGACGACCTGTTACACTAAAGTAGAGTGGGGTGGCTGACCCCTACTTTG	390						
Db	335	AGGTGAGACGACCCACTTACCTAAAGTAGAGTGGGGTGGCTGACCCCTACTTTG	394						
OY	391	GGGCTCCCTGCTGGAAGTCCCTGGCTGAGAGCAGCCAGATCTACACTGTGAGCTTGC	450						
Db	395	GGGCTCCCTGCTGGAAGTCCCTGGCTGAGAGCAGCCAGATCTACACTGTGAGCTTGC	447						
OY	451	AGGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	510						
Db	448	AGGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	506						
OY	511	CAGATGGATCTTCCGATC-AATCCAGATGCGACGACGATGATGATGATGATGATGATG	567						
Db	507	CAGATGGATCTTCCGATC-AATCCAGATGCGACGACGATGATGATGATGATGATGATG	564						

RESULT 13  
AC025574/c  
LOCUS  
DEFINITION Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT  
AC025574  
VERSION AC025574.13 GI:20428723  
KEYWORDS HTG: HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 155023)

Albrooks, S.L., Amaral, H.C., Are, J.R., Ayele, M., Banks, T., Barbieri, J., Benton, J., Blinze, K., Blankenburg, K., Bonin, D., Bouck, J., Bowe, S., Briteva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Dena, A.L., Ding, Y., Dinh, H.H., Donthwaite, K.J., Draper, H., Dugan-Hoch, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabriel, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Jacobson, B., Jia, T., Johnson, R., Jolivet, S., Jume, J., Jackson, L.E., Karlsson, E., Kelly, S., Khan, U., Kling, L., Korah, J., Kovar, C., Kravtsov, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Lounsbury, R., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, T.P., Meador, M., Mei, G., Mettler, M., Moser, M., Neal, D., Newton, J., Mohabbat, K., Morgan, M., Morris, S., Nguyen, N., Nickerson, E., Nwokkenko, S., Ogih, M., Okunou, G., Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pichers, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojiboken, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Tellard, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinsón, R., Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.  
Direct Submission  
Unpublished  
2 (bases 1 to 155023)  
Worley, K.C.  
Direct Submission  
Submitted (11-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 155023)  
Worley, K.C.  
Direct Submission  
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 3, 2002 this sequence version replaced g1:20335511.  
-----  
Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
-----  
Project Information  
Center project name: HALY  
Center clone name: RP11-348M3  
-----  
Summary Statistics  
Sequencing vector: M13  
Chemistry: Dye-primer Body: 5% of reads  
Assembly: Dye-terminator Big Dye: 95% of reads  
Consensus quality: 141522 bases at least Q40  
Consensus quality: 147097 bases at least Q30  
Consensus quality: 150442 bases at least Q20  
Estimated insert size: 150706; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation  
-----

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_diff\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_diff_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
3034 3033: contig of 3033 bp in length  
3134 3033: contig of 3033 bp in length  
5972 3033: contig of 3033 bp in length  
6072 3033: contig of 3033 bp in length  
10270 3033: contig of 3033 bp in length  
10370 3033: contig of 3033 bp in length  
13010 3033: contig of 3033 bp in length  
13110 3033: contig of 3033 bp in length  
17574 3033: contig of 3033 bp in length  
17674 3033: contig of 3033 bp in length  
24651 3033: contig of 3033 bp in length  
24751 3033: contig of 3033 bp in length  
31779 3033: contig of 3033 bp in length  
31879 3033: contig of 3033 bp in length  
42264 3033: contig of 3033 bp in length  
42364 3033: contig of 3033 bp in length  
51105 3033: contig of 3033 bp in length  
51205 3033: contig of 3033 bp in length  
68307 3033: contig of 3033 bp in length  
68407 3033: contig of 3033 bp in length  
109482 3033: contig of 3033 bp in length  
109581 3033: contig of 3033 bp in length



ORIGIN

Query Match 16.6% Score 177.4; DB 2; Length 182350;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-35;  
 Matches 190; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
 Oy 555 TGTGAGGTACTTATGCGCCCTCAGAGCCCTCAGAGTGTGTTGAGAGTGTG 614  
 Db 172034 TTTCTAGGTGCTTATGCGCCCTCAGAGCCCTCAGAGTGTGTTGAGAGTGTG 614  
 Oy 615 TGACCGAATGAAGAGTACGGGGAACAGATGACCTTACACCCAGCAAGTACGT 674  
 Db 171974 TGACCGGATGAAGAGTATGGGGAACAGATGACCTTACACCCAGCAAGTACGT 674  
 Oy 675 ACGGCTGCTGAGCCGGAATGAGAGTACCTGAGTACCTGAGGCAATGCAATGCA 734  
 Db 171914 ACGGTGTGTGCGCCGGAATGAGAGTACCTGAGTACCTGAGGCAATGCA 734  
 Oy 735 CTCAGATATCAGCGCACCTCAAGTTTGGC 765  
 Db 171854 CTCAGATATATAGCGGACCTCAAGTTTGGC 765

COMMENT

Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Glibos, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 184762)  
 Worley, K.C.  
 Direct Submission  
 Submitted (15-FEB-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 184762)  
 Worley, K.C.  
 Direct Submission  
 Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Aug 4, 2002 this sequence version replaced gl:22094227.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Drafting Center Code: BCM  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: HAFc  
 Center clone name: RP11-183H16  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Sequencing vector: M13  
 Chemistry: Dye-Primer Body: 2% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 183845 bases at least Q40  
 Consensus quality: 184849 bases at least Q40  
 Estimated insert size: 185447 bases at least Q20  
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

RESULT 15  
 AC023500 184762 bp DNA linear HTG 24-AUG-2002  
 LOCUS Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
 AC023500  
 AC023500.32 GI:22095070  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 184762)  
 Muzny, D.M., Adams, C., Adlo-Oduola, B., Ali-Osman, F.R., Allen, C.,  
 Albrooke, S.L., Amaralunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
 Barberia, J., Benton, J., Blmage, K., Blankenburg, K., Bonin, D.,  
 Bouck, J., Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Din, H.H.,  
 Doutharte, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Correll, J.H., Guevara, M., Gunatane, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hayes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,  
 Homel, F., Howard, S., Huber, J., Hui, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karisson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kurachi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, Y., Liu, W., Louised, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokwenkwo, S., Oguh, M., Okunolu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Qulles, M., Ren, Y.,  
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
 Scherer, S., Scott, G., Shen, H., Shoostlari, N., Sisson, I.,  
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Swalek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,

FEATURES

1. 184762  
 Location/Qualifiers  
 source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-183H16"

BASE COUNT

47425 a 44751 c 45101 g 46673 t 812 others



GenCore version 5.1.4\_p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 17:35:17 ; Search time 253.709 Seconds  
(without alignments)  
9488.757 Million cell updates/sec

Title: US-10-082-502-18

Perfect score: 1069  
Sequence: 1 gaattcgacagggggggtc.....aaaaaatttcgcgcgcgc 1069

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1069	100.0	1069	20	AAK06970 Mouse secretory pe
2	1069	100.0	1069	21	AAZ08294 Mouse ortholog gen
3	691	64.6	1210	22	AAK21330 Human CDNA sequenc
4	676.4	63.3	1180	20	AAK22112 Human secreted pro
5	606.6	56.7	814	21	AAZ38327 Human transmembran
6	606.6	56.7	814	24	ABK52765 CDNA encoding tran
7	606.6	56.7	814	24	ABU41995 Nucleotide sequenc
8	606.6	56.7	814	24	ABK09772 Human ovarian tumo
9	599.6	56.1	806	20	AAK06969 Human secretory pe

10	599	56.0	806	21	AAZ08293 Human Zs199 gene e
11	564.6	52.8	832	20	AAK97837 Human secreted pro
12	555.2	51.9	714	20	AAK97884 Human secreted pro
13	513.4	48.0	592	21	AAA43525 Mouse secreted exp
14	503.2	47.1	657	24	ABK36007 CDNA sequence #398
15	472.4	44.2	546	21	AAZ38326 Human transmembran
16	345.4	32.3	1085	20	AAK22130 Human secreted pro
17	281.2	26.3	792	20	AAK97836 Human secreted pro
18	256.2	24.0	718	20	AAZ10655 CDNA encoding a hu
19	253	23.7	649	20	AAK97883 Human secretory pe
20	246.8	23.1	649	20	AAK06969 Human secretory pe
21	246.6	23.1	649	21	AAZ08284 Human Zs199 gene.
22	243.4	22.8	558	22	AAZ18023 Human breast cancer
23	238.4	22.3	415	20	AAK06971 Expressed Sequence
24	237.8	22.2	415	21	AAZ08285 Expressed Sequence
25	233.2	21.8	595	22	AAZ07882 Human breast cancer
26	231	20.7	564	24	ABK09691 Human ovarian tumo
27	210.8	19.7	484	22	AAZ25628 Human breast cancer
28	210.8	19.7	1160	22	AAZ12689 Human breast cancer
29	160.2	15.0	352	22	AAK09488 Human secreted pro
30	157.8	14.8	265	20	AAK39549 Human pancreatic c
31	155.6	14.6	2532	22	ABA07341 Human pancreatic c
32	155.6	14.6	2532	22	AAK32768 Human genomic DNA
33	155.6	14.6	2724	22	ABA07340 Human pancreatic c
34	155.6	14.6	2724	22	AAK32767 Human genomic DNA
35	153.6	14.4	495	22	AAI80992 Human polynucleoti
36	84	7.9	215	21	AAC24962 Human secreted pro
37	66.8	6.2	147	24	ABN94837 Gene #1335 used to
38	66.8	6.2	147	24	ABL65469 Lung cancer relate
39	65	6.1	65	24	ABN55235 Gene #1334 used to
40	51.6	4.8	126	24	ABN94836 Lung cancer relate
41	51.6	4.8	126	24	ABL65468 Human spliced tran
42	51	4.8	60	24	ABN39205 Human prostate exp
43	49	4.6	467	23	ABY48839 Human prostate exp
44	48	4.5	475	23	ABY33270 Human prostate exp
45	48	4.5	475	23	ABY42193 Human prostate exp

#### ALIGNMENTS

RESULT 1	AAK06970	AAK06970 standard; cDNA; 1069 BP.
ID	AAK06970;	
XX		
AC	AAK06970;	
XX		
DT	10-MAY-1999	(first entry)
XX		
DE	Mouse secretory peptide-9 (Zs199) orthologue cDNA.	
XX		
KW	Secretory peptide-9; Zs199; orthologue; mouse; tumour marker;	
KM	Cancer; therapy; diagnosis; growth enhancer; ss.	
XX		
OS	Mus sp.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	358..906
FT		/*tag= a
FT	sig_peptide	358..417
FT		/*tag= b
FT	mat_peptide	418..903
FT		/*tag= c
XX		
PN	MO9901554-A1.	
XX		
XX	14-JAN-1999.	
PD		
XX		
PF	02-JUL-1998;	98MO-US13859.
XX		
PR	17-JUN-1998;	98US-0099005.
PR	03-JUL-1997;	98US-0051704.
PR	03-JUL-1997;	97US-0888088.

```

PR 19-MAY-1998; 98US-0081338.
PR 19-MAY-1998; 98US-0083983.
PR 17-JUN-1998; 98US-0089899.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;
XX
XX WPI; 1999-106055/09.
XX
XX P-PSDB; AAM88476.
XX
XX New mammalian secretory peptide-9 (Zsig9) - used as a growth
XX enhancer for placenta, liver and heart, and as an indicator of
XX cancer
XX
XX Claim 2; Page 75-77; 85bp; English.
XX
XX This cDNA clone encodes novel mouse secretory peptide-9, or Zsig9
XX (see AAM88476), an orthologue of novel human Zsig9 (see AAM88469).
XX Human Zsig9 is overexpressed in a number of tumours including
XX brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid
XX and lymphoma tumors, and thus can be used as an indicator for
XX cancer. The invention provides polynucleotides (see AAX06968-70)
XX encoding Zsig9 polypeptides (see AAM88469-77) including mature
XX polypeptides, other processed forms, variants and the mouse
XX orthologues. The Zsig9 gene, or probes derived from it, can be
XX used to determine if Zsig9 is present on chromosome 10, and if a
XX mutation has occurred. Antibodies raised against Zsig9 can be
XX used as diagnostic agents to determine the presence of Zsig9, and
XX thus the presence of cancer. They can also be labelled with
XX radioisotopes or fused with toxins and used to treat tumours
XX which overexpress Zsig9. Antisense nucleotides derived from Zsig9
XX cDNA can also be used to inhibit the growth of tumour cells.
XX Zsig9 proteins can be used to enhance the growth or development of the
XX placenta, heart or liver.
XX
XX Sequence 1069 BP: 276 A; 259 C; 325 G; 209 T; 0 other:
XX
XX Query Match 100.0%; Score 1069; DB 20; Length 1069;
XX Best Local Similarity 100.0%; Pred. No. 8.9e-268;
XX Matches 1069; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 GAATTCGGCAGAGGGGGTCTGCTGCTCCGAGGGGCTCTTAACCTGCTGCTGC 60
XX 1 GAATTCGGCAGAGGGGGTCTGCTGCTCCGAGGGGCTCTTAACCTGCTGCTGC 60
XX
XX 61 GCGAGAGTTTGGAGGGGGGCTTAGGGTTCAGTTTCGGTGGGGGCTGCAGCGGACCT 120
XX 61 GCGAGAGTTTGGAGGGGGGCTTAGGGTTCAGTTTCGGTGGGGGCTGCAGCGGACCT 120
XX
XX 121 CAGATCTCCGCTTAGGTCCTTAGTTAAGTCGGGGAAGCTGGGCGGCTGCTGCGCA 180
XX 121 CAGATCTCCGCTTAGGTCCTTAGTTAAGTCGGGGAAGCTGGGCGGCTGCTGCGCA 180
XX
XX 181 CCGTGAACCTGGCGGGAGCGGAGGCTGAGAGACCGGGACGCCCCGTTTTCCTCA 240
XX 181 CCGTGAACCTGGCGGGAGCGGAGGCTGAGAGACCGGGACGCCCCGTTTTCCTCA 240
XX
XX 241 GCCAGCTCTGAGGGTTGGGAGCCACAGAAACAAAGTGAAGTCCGGCTGCTTCAGAG 300
XX 241 GCCAGCTCTGAGGGTTGGGAGCCACAGAAACAAAGTGAAGTCCGGCTGCTTCAGAG 300
XX
XX 301 CCTGGGCCACGCGCGGCGCCGTCGGAGCAGAGGTGAGCAGCCGTTTACATAAGATG 360
XX 301 CCTGGGCCACGCGCGGCGCCGTCGGAGCAGAGGTGAGCAGCCGTTTACATAAGATG 360
XX
XX 361 AAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGTCCCTGCGGAACCTGGGCTCGA 420
XX 361 AAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGTCCCTGCGGAACCTGGGCTCGA 420
XX
XX 421 AGAGCCCAAGATCTTACACTGTGAGCTTGCAAGGCTCTGGTGATGAATTGAGTGGAA 480
XX 421 AGAGCCCAAGATCTTACACTGTGAGCTTGCAAGGCTCTGGTGATGAATTGAGTGGAA 480

```

```

XX 481 ATTGCCCGGTGGAGCCCAAGAACCATTCAGATGGATCCTTCCGATCAATCCAGAT 540
XX 481 ATTGCCCGGTGGAGCCCAAGAACCATTCAGATGGATCCTTCCGATCAATCCAGAT 540
XX
XX 541 GGCACCCAGTCAAGTTGTGGAGGTACCTTATGCCCCCTCAGAGGCCCTCAGAGATTG 600
XX 541 GGCACCCAGTCAAGTTGTGGAGGTACCTTATGCCCCCTCAGAGGCCCTCAGAGATTG 600
XX
XX 601 CTGAGAGAGTGTGTGACCGAATGAAGAGTACGGGGAAACAGATGACCTTCTACCCAC 660
XX 601 CTGAGAGAGTGTGTGACCGAATGAAGAGTACGGGGAAACAGATGACCTTCTACCCAC 660
XX
XX 661 CGCAGAACTACGATGCGGCTGTGAGCCGGAATGAGAAATCCAGTAAGTACCTTACAG 720
XX 661 CGCAGAACTACGATGCGGCTGTGAGCCGGAATGAGAAATCCAGTAAGTACCTTACAG 720
XX
XX 721 GGCATCCGAAATGATCAGATATTCAGCGGACCTCAAGTTGGCTGTGAGCATTTGTG 780
XX 721 GGCATCCGAAATGATCAGATATTCAGCGGACCTCAAGTTGGCTGTGAGCATTTGTG 780
XX
XX 781 GAGAAATACGAGATGAGCTTATCGAATTTCTCCAGAGAGCTGACAGGTTAAAGAC 840
XX 781 GAGAAATACGAGATGAGCTTATCGAATTTCTCCAGAGAGCTGACAGGTTAAAGAC 840
XX
XX 841 AAACCTTTCAGTAAACGAGACAGATCTATGTGACATGCGCTGCACAGATCTCAGATGAG 900
XX 841 AAACCTTTCAGTAAACGAGACAGATCTATGTGACATGCGCTGCACAGATCTCAGATGAG 900
XX
XX 901 CTAATATCAGTACGAGCAAGACGCTTACACCAAGATGGAACACCCCGAGGGGGA 960
XX 901 CTAATATCAGTACGAGCAAGACGCTTACACCAAGATGGAACACCCCGAGGGGGA 960
XX
XX 961 AGATGCGAGATGCGCTTATATACGTTTATGAAATGAATGAATGAATGAATGAATGA 1020
XX 961 AGATGCGAGATGCGCTTATATACGTTTATGAAATGAATGAATGAATGAATGAATGA 1020
XX
XX 1021 AAACCGAAAGTAAAAAATTTCCGGGCGCC 1069
XX 1021 AAACCGAAAGTAAAAAATTTCCGGGCGCC 1069
XX
XX
XX RESULT 2
XX AAZ08294
XX ID AAZ08294 standard; DNA; 1069 BP.
XX
XX AC AAZ08294;
XX
XX DT 07-FEB-2000 (first entry)
XX
XX DE Mouse ortholog gene encoding Zsig9 secretory protein.
XX
XX KW Secretory protein-9; mouse Zsig9; ortholog; overexpression; antagonist;
XX radio-label; polypeptide toxin; down regulation; diagnostic;
XX therapeutic; probe; cancer; brain; liver; detection; stomach;
XX lymphoma; ds.
XX
XX OS Mus musculus.
XX
XX FH Key Location/Qualifiers
XX FT CDS 358..906
XX FT /tag= a
XX FT /product= "Mouse Zsig9 secretory protein"
XX FT /note= "Overexpressed in tumours"
XX FT sig_peptide 358..417
XX FT /tag= b
XX FT mat_peptide 418..903
XX FT /tag= c
XX FT /label= Mature_Zsig9_protein_variant-4
XX
XX W09960405-A1.

```





22-FEB-2000; 2000MO-US04414.  
 PR 24-FEB-2000; 2000MO-US04914.  
 PR 24-FEB-2000; 2000MO-US05004.  
 PR 01-MAR-2000; 2000MO-US05601.  
 PR 20-MAR-2000; 2000MO-US07377.  
 PR 21-MAR-2000; 2000MO-US07532.  
 PR 30-MAR-2000; 2000MO-US08439.  
 PR 17-MAY-2000; 2000MO-US13705.  
 PR 22-MAY-2000; 2000MO-US14042.  
 PR 30-MAY-2000; 2000MO-US14941.  
 PR 02-JUN-2000; 2000MO-US15264.  
 PR 10-NOV-2000; 2000MO-US30873.  
 XX (GEM) GENENTECH INC.  
 PA Baker RP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WT, Zhang Z;  
 XX WPI: 2001-408281/43.  
 DR P-PSDB: AA012258.  
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX Claim 3; Fig 173; 813pp; English.  
 PS AAS1244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumor necrosis factor alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, or the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 CC XX  
 XX Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;  
 SQ  
 Query Match 64.6%; Score 691; DB 22; Length 1210;  
 Best Local Similarity 84.4%; Pred. No. 1.6e-169;  
 Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

DB 435 -TAGTTTCGGAGGCCACTGGGGCCAAAGTGAAGTCCACGGCTCTCCAGCGCTTGGGCC 493  
 OY ACCGCGCGCGCCCTGGAGACAGAGTGGAGCGACCCCTGTTTACACTAAGATGAAGGCTG 368  
 DB 494 ACGGCGCGCGCCCTGGAGACAGAGTGGAGCGACCCCTTACGCTTAAATGAAGGCTG 553  
 OY 369 GGGTTGGCTAGCCCTACTTTTGGGGGCTCTGCTGGGACCTGGCTGGAGGAGGCA 428  
 DB 554 GGGTTGGCTAGCCCTCTCTTGGGGGCTCTGCTGGGAGAACCGCTGGGCTCGAGAGGCA 613  
 OY 429 AGATCTACACTGTGGAGCTTGGAGGCTCTGGTGGATGAATTAGATGGGAATTCGCCG 488  
 DB 614 GGAATCTCCACTGTGGAGCATGACAGGGCTCTGTGTGAATGAATGGGAATTCGCCA 673  
 OY 489 CGTGACCCCAAGAGACCATTCAGATGGATCTCCCAATCAATCAGATGGAGCA 548  
 DB 674 GGTGACCCCAAGAGACCATTCAGATGGATCTTCGGATCAATCAGATGGAGCA 733  
 OY 549 GTGAGTTGTGAGGTACCTTATGCCGCTCAGAGGCCACCTCAGAGTTCCTTGAGGA 608  
 DB 734 GTGAGTTGTGAGGTACCTTATGCCGCTCAGAGGCCACCTCAGAGTTCCTTGAGGA 793  
 OY 609 GGTGTGTACCGAATGAAGAGTACGGGAGAACAGATTCCTTACCCAGGAGCA 668  
 DB 794 GATATGTGACCGAATGAAGAGTACGGGAGAACAGATTCCTTACCCAGGAGCA 853  
 OY 669 CTACGTACCGCTGTGAGGAGCGGGAATGAGATTCAGTAACTTACAGAGGATCCG 728  
 DB 854 CTACGTACCGCTGTGAGGAGCGGGAATGAGATTCAGTAACTTACAGAGGATCCG 913  
 OY 729 AATGTACCTAGATATCAGGAGGAGCCCTCAAGTTGGCTGAGAGATTTGGAAGATA 788  
 DB 914 AATGTACCTAGATATCAGGAGGAGCCCTCAAGTTGGCTGAGAGATTTGGAAGATA 973  
 OY 789 CGAGATGAGCTTTCGATTTCTTCCAGAGAGGTGCAACGTTAAAGAACTTGG 848  
 DB 974 CGAGATGAGCTTTCGATTTCTTCCAGAGAGGTGCAACGTTAAAGAACTTGG 1033  
 OY 849 CAGTAAGCGAGAGATCTATGAGCATGCTCCGACAGATTCAGATGATGATGAT 908  
 DB 1034 CAGTAAGCGAGAGATCTTGTGACATGCTCCGACATATGCAATGATGATGATGAT 1093  
 OY 909 CACTGAGCAGACGCTTACACCAACGATGAGAAACCCCGAGAGGAGATGCA 968  
 DB 1094 CACTGAGCAGACGCTTACACCAACGATGAGAAACCCCGAGAGGAGATGCA 1148  
 OY 969 GCATGCGCTT--TATATGCTTTTATGAAATGAACTGAAAAA 1013  
 DB 1149 GCATGCGCTTATATATGATTTTACTGAAATTAATGAAAAA 1195  
 DE Human secreted protein gene 2 clone H2MB56.  
 XX Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
 XX tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
 XX developmental abnormality; foetal deficiency; Alzheimer's disease;  
 XX cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
 XX immune deficiency disease; respiratory disorder; arthritis; skeletal;  
 XX haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
 XX cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
 OS Homo sapiens.  
 XX W09901020-A2.



PD 04-NOV-1999.  
 XX 27-APR-1999; 99WO-JP02226.  
 PF 28-APR-1998; 98JP-0119395.  
 PR (SAGA) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX Kato S, Kimura T;  
 PI WPI: 2000-023358/02.  
 DR P-PSDB: AAY52391.  
 PT Human proteins with transmembrane domains, involved in control of cell  
 PT proliferation and differentiation, useful for treating e.g. cancer or  
 PT inflammation  
 XX  
 XX  
 PS Claim 4: Page 106-107; 114pp; English.  
 CC This sequence represents the human cDNA clone HP10390  
 CC which encodes a 20 kb protein with one putative transmembrane  
 CC domain in the N-terminus. The cDNA was isolated from a human stomach  
 CC cancer cell line cDNA library. The protein has no homology with any  
 CC known protein. The protein may be used to raise specific antibodies, as  
 CC assay reagents, as diagnostic tissue markers, for the isolation of  
 CC cognate receptors, ligands and binding proteins, and as biologically  
 CC active agents. Nucleotides encoding the protein may be used as primers  
 CC and probes or antisense molecules, and in gene therapy. Cells transfected  
 CC with these nucleotides may be used to screen for agonists and antagonists  
 CC which are potentially useful therapeutically.  
 XX  
 SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
 Query Match 56.7%; Score 606.6; DB 21; Length 814;  
 Best Local Similarity 87.3%; Pred. No. 1.2e-147;  
 Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;  
 OY 213 AGAAGCCGGGAGAGCCCGCTTTTCCAGCCAGCTGCTAGGGTTGGGACCCACAGAAAC 272  
 DB 1 ACAAATCCCGAGAGCCCTGCTCCCTGACGACGAGTG-7AGTTTCGGAGCCACTGGGCC 59  
 OY 273 AAGTGAAGTCCGGCTGCTTCCAGAGCCTGGCCAGCGGGCGGCGCTGGGAGCAGAG 332  
 DB 60 AAGTGAAGTCCAGCGGCTTCCAGCGCTTGGCCAGCGGGCGGCGCTGGGAGCAGAG 119  
 OY 333 GTGAGCAGACCTGTTACATAAAGATGAAGAGTGGGGTGGCTAGCCCTATTTTGGG 392  
 DB 120 GTGAGCAGACCCATTATAGCTAAAGATGAAGAGTGGGGTGGCTAGCCCTATTTGGG 179  
 OY 393 GGTCTGCTGGGAACTGCTGGGCTCGAAGAGCCAAATCTACACTGTGGAGCTTGCAG 452  
 DB 180 GGCCTGCTGGGAACTGCTGGGCTCGAAGAGCCAAATCTACACTGTGGAGCTTGCAG 239  
 OY 453 GGCCTGCTGGATGAATTAAGTGGGAAATTCGCCGCGTGGAGCCCAAGAAACCATTA 512  
 DB 240 GGCCTGCTGGATGAATTAAGTGGGAAATTCGCCGCGTGGAGCCCAAGAAACCATTA 299  
 OY 513 GATGGGATCTTCCGAATCAATCGAATGGCAGGCACTGATGTTGGAGTACTTATGC 572  
 DB 300 GATGGGATCTTCCGGAATCAATCGAATGGCAGGCACTGATGTTGGAGTACTTATGC 359  
 OY 573 CCGCTCAGAGGCCCACTTCACAGAGTGTGGAGAGGTGTGTGACCGAATGAGAGTA 632  
 DB 360 CCGCTCAGAGGCCCACTTCACAGAGTGTGGAGAGGTGTGTGACCGAATGAGAGTA 419  
 OY 633 CGGGAGACAGATTGACCTTACCCACCGAAGAACTACGATGCGTGTGAGCGGGA 692  
 DB 420 TGGGAGAAAGATTGATCTTCCACCCATGCGAAGAACTACGATGCGTGTGAGCGGGA 479  
 OY 693 TGGAGAAATCGAGTAACTAGACTTACAGAGGCAATCGAATTTGATGATATGAGCGGCAC 752  
 DB 480 TGGAGAAATCGAGTAACTAGAGTAACTAGAGGCAATCGAATTTGATGATATGAGCGGCAC 539

OY 753 CCTCAAGTTTGGCTGTGAGAGCATTTGTGGAAGAAATACGAGATGACCTTATGCAATTC 812  
 DB 540 CCTCAAGTTTGGCTGTGAGAGCATTTGTGGAAGAAATACGAGATGACCTTATGCAATTC 599  
 OY 813 CTCACAGAGGCTGACACAGTTTAAAGCAAACTTTCAGTAAAGCGGACAGATCTAGTGA 872  
 DB 600 TTCACAGAGGCTGACACAGTTTAAAGCAAACTTTCAGTAAAGCGGACAGATCTTGTGA 659  
 OY 873 CCATGCCCTGACAGATCTACAGATGAGTATGAAATCACTGAGCAAGAGCCTACACCA 932  
 DB 660 CCATGCCCTGACAGATCTACAGATGAGTATGAAATCACTGAGCAAGAGCCTACACCA 714  
 OY 933 AACGTATGGAACCCCGACAGAGGGAAGATGACAGATTCCTT--TATATTACGTT 990  
 DB 715 GCGTTATGATGATCACCACCGACAGAGGGAAGATGATGCAATGCTTTATATATGTT 774  
 OY 991 TTTATGGAATGAACTGAAAAA 1013  
 DB 775 TTTACTGAAATTAATGAAAAA 797  
 RESULT 6  
 ABK52765  
 ID ABK52765 standard; cDNA; 814 BP.  
 XX  
 XX ABK52765;  
 AC  
 XX  
 XX 15-AUG-2002 (first entry)  
 DT  
 XX  
 XX cDNA encoding transmembrane protein 4, a cancer-linked protein.  
 DE  
 XX  
 KW Expressed sequence tag; EST; human; cancer; anti-neoplastic;  
 KW cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;  
 KW transmembrane protein 4; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 145..693  
 FT /tag= a  
 FT /product= "Transmembrane protein 4"  
 XX  
 PD WO200231198-A2.  
 XX  
 PD 18-APR-2002.  
 PF 11-OCT-2001; 2001WO-US31607.  
 XX  
 PR 11-OCT-2000; 2000US-239294P.  
 PR 11-OCT-2000; 2000US-239297P.  
 PR 11-OCT-2000; 2000US-239605P.  
 PR 12-OCT-2000; 2000US-239802P.  
 PR 12-OCT-2000; 2000US-239805P.  
 PR 12-OCT-2000; 2000US-239806P.  
 PR 16-OCT-2000; 2000US-240622P.  
 PR 19-OCT-2000; 2000US-241682P.  
 PR 19-OCT-2000; 2000US-241723P.  
 PR 31-OCT-2000; 2000US-244932P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Horrigan S, Weaver Z, Endress GA;  
 DR WPI: 2002-463271/49.  
 DR P-PSDB: AAV97063.  
 XX  
 PT Identifying modulators of a cancer-related gene to screen agents for  
 PT preventing or treating cancer comprises detecting a difference in the  
 PT expression of cancer-linked genes in the presence or absence of test  
 PT compounds -  
 PS Claim 1: Page 43; 66pp; English.

XX The invention relates to modulators of a cancer-related genes. Also  
 CC described are: (1) processes for identifying an anti-neoplastic agents  
 CC comprising contacting a cell exhibiting neoplastic activity with a  
 CC compound first identified as a cancer related gene modulator, and  
 CC detecting a decrease in the neoplastic activity. (2) a process for  
 CC determining the cancerous state of a cell by determining an increase in  
 CC the level of expression of at least one gene, where an elevated  
 CC expression relative to a known non-cancerous cell indicates a cancerous  
 CC state or potentially cancerous state. The anti-neoplastic agent is  
 CC useful for treating cancer or for protecting an animal against cancer.  
 CC The immunogenic composition is also useful for treating cancer in an  
 CC animal, where the composition elicits the production of cytotoxic T  
 CC lymphocytes specific for the immunogenic composition. Preferably, the  
 CC animal is a human. The cancer-linked genes and polypeptides are also  
 CC useful as targets for cancer therapy or chemotherapy. The present  
 CC sequence represents a cancer-linked gene located on chromosome 12,  
 CC which encodes transmembrane protein 4.

SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match 56.7%; Score 606.6; DB 24; Length 814;

Best Local Similarity 87.3%; Pred. No. 1.2e-147;

Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;

OY 213 AGAAGCCGGGACAGCCCGTTTCCAGCCAGCTGCTAGGGTTGGAGCCACAGAAAC 272  
 DB 1 AGAATCCCGGACAGCCCGTCTCCAGCCAGCTGCTAGGGTTGGAGCCAGTGGAGCC 59  
 OY 273 AAATGAGAGTCCGGGCTTTCCAGAGCCTGGGGCCAGCGGGCGGCGGCGGGAGAGCAG 332  
 DB 60 AAATGAGAGTCCAGGCTCTCCAGGCGCTGGCCAGCGGGCGGCGGCGGGAGAGCAG 119  
 OY 333 GTGAGAGGACCTTACATTAAGATGAAGAGCTGGGGTTGGCTAGCCCTACTTTTGGG 392  
 DB 120 GTGAGAGGACCTTACATTAAGATGAAGAGCTGGGGTTGGCTAGCCCTACTTTTGGG 179  
 OY 393 GGTCTGCTGGGAACTGCTGGGCTGAGAGGACCAAGATCTACATCTGGAGCTTGACAG 452  
 DB 180 GGGCTGCTGGGAACTGCTGGGCTGAGAGGACCAAGATCTACATCTGGAGCTTGACAG 239  
 OY 453 GGGCTGCTGGGAACTTAAAGTGGGAAATTTGCCCGGTGGAGCCCAAGAAACCATATCA 512  
 DB 240 GGGCTGCTGGGAACTTAAAGTGGGAAATTTGCCCGGTGGAGCCCAAGAAACCATATCA 299  
 OY 513 GATGGATCTCTCCGATCAATCCAGATGGAGCAGTCAAGTGTGGAGGTTACTTATGC 572  
 DB 300 GATGGATCTCTCCGATCAATCCAGATGGAGCAGTCAAGTGTGGAGGTTACTTATGC 359  
 OY 573 CCGCTCAGAGGCCACCTCAGAGAGTGTGTTGAGAGGTTGTGACCCGAATGAAGAGTA 632  
 DB 360 CCGCTCAGAGGCCACCTCAGAGAGTGTGTTGAGAGGTTGTGACCCGAATGAAGAGTA 419  
 OY 633 CGGGAACAGATGACCTCTACCCACCGCAAGAACTAGCTACGCTCTGTGAGCCGGA 692  
 DB 420 TGGGGAACAGATGATCTTCCACCCCTCCAAAGAACTAGCTAGTGTGAGCCGGA 479  
 OY 693 TGGAGATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 752  
 DB 480 TGGAGATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539  
 OY 753 CTTCAAGTTTGGCTGAGAGCATTTGAGAGAAATACGAGATGAGCTTATTCGAATCTT 812  
 DB 540 CTTCAAGTTTGGCTGAGAGCATTTGAGAGAAATACGAGATGAGCTTATTCGAATCTT 599  
 OY 813 CTTCAAGAGGCTGAGAGCATTTAAAGCAAACTTTGAGAGAAACGAGCAATCTATGTA 872  
 DB 600 TTTCCGAGAGGCTGAGAGCATTTAAAGCAAACTTTGAGAGAAACGAGCAATCTTGTGTA 659  
 OY 873 CCAATGCCCTGACAGATCTGACAGTATGATGATGATGATGATGATGATGATGATGATGAT 932  
 DB 660 CCAATGCCCTGACAGATCTGACAGTATGATGATGATGATGATGATGATGATGATGATGAT 714

OY 933 AACGTGATGGAACACCCCGAGAGGAGAGATGCGACATGCTTT--TATATTACGTT 990  
 DB 715 GGGTTATGATGATACCCCGAGAGGAGAAATGTTGGCATGCTTTATATTATGTT 774  
 OY 991 TTTATGGAATGAGCTGAAAAA 1013  
 DB 775 TTTACTGAATTTACTGAAAAA 797

# RESULT 7

ID ABL41995 standard; DNA; 814 BP.

AC ABL41995;

DT 11-JUN-2002 (first entry)

DE Nucleotide sequence of human polypeptide HP10390.

KW Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.

XX Homo sapiens.

FT Key Location/Qualifiers

FT CDS 145..693 /tag- a /product- "HP10390"

PN W0200208416-A1.

PD 31-JAN-2002.

PE 24-JUL-2001; 2001WO-JP06371.

PR 24-JUL-2000; 2000JP-0222743.

PR 24-AUG-2000; 2000JP-0254407.

PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;

DR WPI; 2002-195877/25.

PI P-PSDB; ABB09717.

PT Antibody preparation by inoculation of an animal with a vector

PT expressing a fusion protein of an antigen on the C-terminal side of a

PT transmembrane domain for use as drugs, diagnostic reagents and

PT laboratory reagents

XX Example; Page 33-35; 45pp; Japanese.

PS The specification describes a method of antibody preparation. The

CC method comprises inoculating an animal with a vector expressing

CC a fusion protein having an antigen protein fused to the C-terminal

CC side (extracellular) of a transmembrane domain protein (the

CC N-terminal side of which is intracellular), and then isolating and

CC purifying the antibody from the animal. The antibodies can be used

CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The

CC present sequence encodes a polypeptide, designated HP10390, which was

CC used in the course of the invention.

SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Query Match 56.7%; Score 606.6; DB 24; Length 814;

Best Local Similarity 87.3%; Pred. No. 1.2e-147;

Matches 701; Conservative 0; Mismatches 94; Indels 8; Gaps 3;

OY 213 AGAAGCCGGGACAGCCCGTTTCCAGCCAGCTGCTAGGGTTGGAGCCACAGAAAC 272.  
 DB 1 AGAATCCCGGACAGCCCGTCTCCAGCCAGCTGCTAGGGTTGGAGCCACAGTGGGCC 59  
 OY 273 AAATGAGAGTCCGGGCTTTCCAGAGCCTGGGCCAGCGGGCGGCGGGAGAGCAGAC 332  
 DB 1 AGAATCCCGGACAGCCCGTCTCCAGCCAGCTGCTAGGGTTGGAGCCACAGTGGGCC 59





QY 693 TGGAGAAATCCAGTACTAGACTTACAGGCGATCCGAATTGACTCAGATATACGCGCAC 752  
 Db 480 TGGAGAAATCCAGTACTAGACTTACAGGCGATCCGAATTGACTCAGATATACGCGCAC 559  
 QY 753 CCTCAAGTTTCCGTGAGAGCATTTGGAAGAATAGAGGATGAGCTTATCGAATCTT 812  
 Db 540 CCTCAAGTTTCCGTGAGAGCATTTGGAAGAATAGAGGATGAGCTTATCGAATCTT 599  
 QY 813 CTCCAGAGAGGCTGACACCTTAAAGACAACCTTTCAGTACGAGAGATCTATGTGA 872  
 Db 600 TTCCCGAGAGGCTGACACCTTAAAGACAACCTTTCAGTACGAGAGATCTATGTGA 659  
 QY 873 CCATGCCCTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 932  
 Db 660 CCATGCCCTCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 714  
 QY 933 AACGTATGAGACACCCCGAGAGGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 990  
 Db 715 GGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774  
 QY 991 TTATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013  
 Db 775 TTATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 797

RESULT 9  
 AAX06969  
 ID AAX06969 standard; cDNA; 806 BP.  
 AC AAX06969;  
 XX  
 XX  
 DT 10-MAY-1999 (first entry)  
 XX  
 DE Human secretory peptide-9 (Zs19) variant cDNA.  
 XX  
 KW Secretory peptide-9; Zs19; human; tumour marker; cancer; therapy;  
 XX diagnosis; growth enhancer; variant; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 104..652  
 FT /\*tag- a  
 FT sig-peptide 104..163  
 FT /\*tag- b  
 FT mat-peptide 164..649  
 FT /\*tag- c  
 XX  
 PN WO9901554-A1.  
 XX  
 PD 14-JAN-1999.  
 XX  
 PF 02-JUL-1998; 98MO-US13859.  
 XX  
 PR 17-JUN-1998; 98US-0099005.  
 PR 03-JUL-1997; 97US-0051704.  
 PR 13-JUL-1997; 97US-0888088.  
 PR 19-MAY-1998; 98US-0081338.  
 PR 19-MAY-1998; 98US-0083883.  
 PR 17-JUN-1998; 98US-0089899.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TR;  
 DR WPI: 1999-106055/09.  
 DR P-PSDB; AAM88474.  
 XX  
 PT New mammalian secretory peptide-9 (Zs19) - used as a growth  
 PT enhancer for placenta, liver and heart, and as an indicator of  
 PT cancer  
 PS  
 Claim 2: Page 73-74; 85pp; English.

XX This cDNA clone encodes human secretory peptide-9, or Zs19,  
 CC variant (see AAM88474). Zs19 (see also AAM88469) is overexpressed in  
 CC human brain, liver, lung, oesophageal, stomach, colon, rectal,  
 CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an  
 CC indicator for cancer. Zs19 cDNA was discovered in a placenta  
 CC clone from a full-term pregnancy cDNA library which contained an  
 CC expressed sequence tag (see AAX06971). The invention provides an  
 CC polynucleotide (see AAX06968-70) encoding Zs19 polypeptides (see  
 CC AAM88469-77) including mature polypeptides, other processed forms,  
 CC variants and mouse orthologues. The Zs19 gene, or probes derived  
 CC from it, can be used to determine if Zs19 is present on chromosome  
 CC 12, and if a mutation has occurred. Antibodies raised against  
 CC Zs19 can be used as diagnostic agents to determine the presence of  
 CC with radioisotopes or fused with toxins and used to treat tumors  
 CC which overexpress Zs19. Antisense nucleotides derived from Zs19  
 CC cDNA can also be used to inhibit the growth of tumour cells. Zs19  
 CC proteins can be used to enhance the growth or development of the  
 CC placenta, heart or liver.

Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;

Query Match 56.1%; Score 599.6; DB 20; Length 806;  
 Best Local Similarity 87.7%; Pred. No. 7.7e-146;  
 Matches 690; Conservative 1; Mismatches 86; Indels 10; Gaps 3;

QY 272 CAAAGTGAAGTCCGGCTCTTTCAGAGCCCTGGCCAGCGCGCGCGGTGGAGCAGA 331  
 Db 18 CAAAGTGAAGTCCGGCTCTTTCAGAGCCCTGGCCAGCGCGCGCGGTGGAGCAGA 77  
 QY 332 GGTGAGAGCCCGCTTACACTAAGATGAAGGCTGGGTGGCTTACCTTCTTGG 391  
 Db 78 GGTGAGAGCCCGCTTACACTAAGATGAAGGCTGGGTGGCTTACCTTCTTGG 137  
 QY 392 GGGTCTGCTGGGAACTGCGTGGCTGGAAGAGCCCAATCTACCTGNGAGCTTGA 451  
 Db 138 GGGCTCTGCTGGGAACTGCGTGGCTGGAAGAGCCCAATCTACCTGNGAGCTTGA 197  
 QY 452 GGGCTCTGCTGGGAACTGCGTGGCTGGAAGAGCCCAATCTACCTGNGAGCTTGA 511  
 Db 198 GGGCTCTGCTGGGAACTGCGTGGCTGGAAGAGCCCAATCTACCTGNGAGCTTGA 257  
 QY 512 AGATGGATCTCTCCGAATCAATCCAGATGCGAGCCAGTCACTGGAGATACCTTATG 571  
 Db 258 AGATGGATCTCTCCGAATCAATCCAGATGCGAGCCAGTCACTGGAGATACCTTATG 317  
 QY 572 CCGGCTCAAGAGCCCACTCAGAGATGCTTGGAGAGGTGTGTACCGAATGAAGAGT 631  
 Db 318 CCGGCTCAAGAGCCCACTCAGAGATGCTTGGAGAGGTGTGTACCGAATGAAGAGT 377  
 QY 632 ACGGGACAGATTGACCTTTCACCCAGCAAGAACTACGATGCGTGGAGCCGA 691  
 Db 378 ATGGGACAGATTGACCTTTCACCCAGCAAGAACTACGATGCGTGGAGCCGA 437  
 QY 692 ATGGGACAGATTGACCTTTCACCCAGCAAGAACTACGATGCGTGGAGCCGA 751  
 Db 438 ATGGGACAGATTGACCTTTCACCCAGCAAGAACTACGATGCGTGGAGCCGA 497  
 QY 752 CCTCAAGTTTCCGTGAGAGCATTTGGAAGAATAGAGGATGAGCTTATCGAATCT 811  
 Db 498 CCTCAAGTTTCCGTGAGAGCATTTGGAAGAATAGAGGATGAGCTTATCGAATCT 557  
 QY 812 TCTCCAGAGAGGCTGACACCTTAAAGACAACCTTTCAGTACGAGAGATCTATGTG 871  
 Db 558 TCTCCAGAGAGGCTGACACCTTAAAGACAACCTTTCAGTACGAGAGATCTATGTG 617  
 QY 872 AACATGCCCTGAGAGCTCAGATGAGCTATGATGATGATGATGATGATGATGATGATGAT 931  
 Db 618 AACATGCCCTGAGAGCTCAGATGAGCTATGATGATGATGATGATGATGATGATGATGAT 672  
 QY 932 AACATGCCCTGAGAGCTCAGATGAGCTATGATGATGATGATGATGATGATGATGATGAT 989



```
DB 673 TGCCTTGATGCATCACCCCGAGAGGGGAAATGCTGGCATGCCCTTTATATATATCT 732
OY 990 TTTTATGGAATGAACTGAAAAAACTCTGAAACCGAAAGTAAAAA 1049
DB 733 TTTTACTGAAATTAAGTAAAAA---TATGAACCAAAAGTAAAAA 789
OY 1050 AAAAAA 1056
DB 790 AGAGAGA 796

RESULT 10
AAZ08293 standard; DNA; 806 BP.
AAZ08293:
07-FEB-2000 (first entry)
XX Human Zs19 gene encoding secretory protein variant-4.
DE
XX Secretory protein-9; Human Zs19; chromosome 12q15 region; variant;
XX overexpression; antagonist; antibody; antisense nucleotide; tumour;
XX treatment; receptor; radio-label; fusion; polypeptide toxin; technique;
XX down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;
XX detection; stomach; lymphoma; alternative splicing; allelic variation;
XX silent mutation; ds.
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 104..652
XX FT /*tag- A
XX FT /product= "Zs19 secretory protein variant-4"
XX FT /note= "Overexpressed in tumours"
XX FT sig_peptide 104..163
XX FT /*tag- b
XX FT mat_peptide 164..649
XX FT /*tag- C
XX FT /label= Mature_Zs19_protein_variant-4
XX PN W09960405-A1.
XX PD 25-NOV-1999.
XX PE 19-MAY-1999; 99WO-US11107.
XX PR 19-MAY-1998; 98US-0081183.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Moore EE, Taft DW;
XX DR WPI: 2000-039447/03.
XX DR P-PSDB: AAY15135.
XX
XX Detecting tumors using antibodies, antagonists and antisense
XX nucleotides to secretory protein-9 (Zs19)
XX
XX Disclosure; Page 33-35; 45pp; English.
XX
XX The present DNA sequence is a gene encoding the variant of the secretory
XX protein-9, Zs19 that arises due to alternative splicing; allelic
XX variation or silent mutations that result in amino acid changes. This
XX sequence is mapped to the human chromosome 12q15 region. It is
XX overexpressed in tumours. Antagonists, antibodies and antisense
XX nucleotides to Zs19 are useful for detecting and treating tumours. The
XX antagonist may be an antibody or receptor to Zs19 and it may be radio-
XX labelled or fused to a polypeptide toxin. It can be used for down
XX regulating the overexpression of Zs19. The gene sequence can be used as
XX nucleic acid probes to detect RNA encoding Zs19. The Zs19 sequence
XX facilitates improved diagnostic and therapeutic techniques for detecting
```

```
CC and treating cancers, especially of the brain, liver, stomach, lymphoma,
CC etc., at an early stage.
XX
XX Sequence 806 BP; 229 A; 183 C; 233 G; 160 T; 1 other;
SQ
Query Match 56.0%; Score 599; DB 21; Length 806;
Best Local Similarity 87.7%; Pred. No. 1,1e-145;
Matches 690; Conservative 1; Mismatches 86; Indels 10; Gaps 3;
OY 272 CAAAGTGAAGTCCGGCTCTTCCAGAGCTGGGGCCAGCGGGCGCTGGAGACAGA 331
DB 18 CAAAGTGAAGTCCAGCGGTCTGCGACGCTGGCCAGCGGGCGCTGGAGACAGA 77
OY 332 GGTGAGCAGCCCTTTACACTAAAGATGAAAGGCTGGGTTGGCTACCTTATG 391
DB 78 GGTGAGCAGCCCGCTTACCTTAABATGAAAGGCTGGGTTGGCTGGCCCTGCTG 137
OY 392 GGGTCTGCTGGGAAGTGGCTGGCTGGAAGGAGCCAAATCTACACTGTGAGCTTGA 451
DB 138 GGGCCCTGCTGGGAACCGGCTGGGCTGGAGAGCCAGCATCTCCACTGTGAGCATGA 197
OY 452 GGGCTCTGCTGATGATTAAGATGGGAATTCGCCGCTGGACCCCAAGAACCATTC 511
DB 198 GGGCTCTGCTGATGATTAAGATGGGAATTCGCCAGGTGGACCCCAAGAACCATTC 257
OY 512 AGATGGGATCTTCGGAATCAATCCAGATGCGAGCCAGTCAGTTGTGAGTACCTTATG 571
DB 258 AGATGGGATCTTCGGAATCAATCCAGATGCGAGCCAGTCAGTTGTGAGTACCTTATG 317
OY 572 CCCGCTCAGAGCCCGACCTCAGAGAGTTGCTTGAGAGAGTGTGACCAATGAGAGACT 631
DB 318 CCCGCTCAGAGCCCGACCTCAGAGAGTTGCTTGAGAGAGTGTGACCAATGAGAGACT 377
OY 632 ACGGGAACAGATTGACCTTCTACCCAGCAAGAACATGACGATGACGCTGGAGCGGA 691
DB 378 ACGGGAACAGATTGATCTTCCACCATTCGCAAGACTACCTGATGAGTGGCGGA 437
OY 692 ATGGAATCCAGTGAATGACTTACAGAGGATCCGAATTGACTCAGATATCAGCGGA 751
DB 438 ATGGAATCCAGTGAATGACTTACAGAGGATCCGAATTGACTCAGATATCAGCGGA 497
OY 752 CCCGCTCAGAGCCCGACCTCAGAGAGTTGCTTGAGAGAGTGTGACCAATGAGAGACT 811
DB 498 CCCGCTCAGAGCCCGACCTCAGAGAGTTGCTTGAGAGAGTGTGACCAATGAGAGACT 557
OY 812 TCTCCAGAGAGGCTGACAACTTAAAGCAAACTTTCAGTAAAGGAGCAGATCTATGTG 871
DB 558 TTTCCGAGAGGCTGACAACTTAAAGCAAACTTTCAGTAAAGGAGCAGATCTATGTG 617
OY 872 ACCATGCCCTGCACAGATCTCAGATGAGCTATGAATCAGTGCAGACCAAGCCTACACC 931
DB 618 ACCATGCCCTGCACATATCGCATGTGATGAGCTATGAATCAGTGCAGACCAAGCCTACACC 672
OY 932 AAACGTGATGCAACACCCCGAGAGGGGGAAGTGGCAGATTCCTTT--TATATTACGT 989
DB 673 TGGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
OY 990 TTTTATGGAATGAACTGAAAAAACTCTGAAACCGAAAGTAAAAA 1049
DB 733 TTTTACTGAAATTAAGTAAAAA---TATGAACCAAAAGTAAAAA 789
OY 1050 AAAAAA 1056
DB 790 AGAGAGA 796

RESULT 11
AAZ97837 standard; CDNA; 832 BP.
AAZ97837:
XX
XX AAZ97837:
XX 23-SEP-1999 (first entry)
```







Db 596 CAATCCCTTTATATATATGTTTACTGAAATTAACGAAAAA 641

RESULT 15

AAZ38326 ID AAZ38326 standard; cDNA: 546 BP.

AAZ38326; AC

09-FEB-2000 (first entry)

Human transmembrane protein cDNA clone HP10390 coding sequence.

HP10390; transmembrane domain; stomach cancer cell; antibody;

assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;

agonist; antagonist; ligand; therapeutic; ds.

Homo sapiens.

Key Location/Qualifiers

FT 1..540 /tag= a /product= "Human transmembrane protein HP10390"

FT /note= "No stop codon given in the specification"

PN W09955862-A2.

PD 04-NOV-1999.

PE 27-APR-1999; 99WO-JP02226.

PR 28-APR-1998; 98JP-0119395.

PA (SAGA ) SAGAMI CHEM RES CENT.

PI (PROT-) PROTEGENE INC.

PI Kato S, Kimura T;

DR WPI: 2000-023358/02.

DR P-PSDB; AA523391.

XX Human proteins with transmembrane domains, involved in control of cell proliferation and differentiation, useful for treating e.g. cancer or inflammation

XX Claim 3; Page 88; 114pp; English.

CC This sequence represents the coding sequence of human cDNA clone

CC HP10390 which encodes a 20 kD protein with one putative transmembrane

CC domain in the N-terminus. The cDNA was isolated from a human stomach

CC cancer cell line cDNA library. The protein has no homology with any

CC known protein. The protein may be used to raise specific antibodies, as

CC assay reagents, as diagnostic tissue markers, for the isolation of

CC cognate receptors, ligands and binding proteins, and as biologically

CC active agents. Nucleotides encoding the protein may be used as primers

CC and probes or antisense molecules, and in gene therapy. Cells transfected

CC with these nucleotides may be used to screen for agonists and antagonists

CC which are potentially useful therapeutically.

SO Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;

Query Match 44.2%; Score 472.4; DB 21; Length 546;  
Best Local Similarity 91.6%; Pred. No. 7.5e-113;  
Matches 500; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 358 ATGAAGGCTGGGCTGCTAGCCCTACTTTGGGGGCTCTGGAAGTCTGCGCT 417  
DB 1 ATGAAGGCTGGGCTGCTAGCCCTACTTTGGGGGCTCTGGAAGTCTGCGCT 60

QY 418 CGAAGGACCAAGACTCTGAGCTTGAGAGCTTGAGAGCTTGAGAGTGAATGAGTGG 477  
DB 61 CGAAGGACCAAGACTCTGAGCTTGAGAGCTTGAGAGCTTGAGAGTGAATGAGTGG 120

QY 478 GAAATTCGCCGCGTGGAGCCCAAGAACATTCAGATGGGATCTTCGAAATCAATCCA 537  
DB 121 GAAATTCGCCGCGTGGAGCCCAAGAACATTCAGATGGGATCTTCGAAATCAATCCA 180  
QY 538 GATGGCAGCCAGTCACTTTGTGGAGTACCTTTATGCCCGCTCAGAGGCCCACTCAGAGAG 597  
DB 181 GATGGCAGCCAGTCACTTTGTGGAGTACCTTTATGCCCGCTCAGAGGCCCACTCAGAGAG 240  
QY 598 TTGCTTGAGAGAGTGTGTGACCGAATGAAGAGTACGGGAGACAGATTGACCTTAC 657  
DB 241 TTGCTTGAGAGAGTGTGTGACCGAATGAAGAGTACGGGAGACAGATTGACCTTAC 300  
QY 658 CACCGCAAGACTACGTACGCGTGTGACCGGCAATGAGAAATCAAGTGAAGTGA 717  
DB 301 CACCGCAAGACTACGTACGCGTGTGACCGGCAATGAGAAATCAAGTGAAGTGA 360  
QY 718 CAGGCGATCCGAATTCAGTACAGATATCAGCGGACCCCTCAAGTTTGGCTGTGAGACATT 777  
DB 361 CAGGCGATCCGAATTCAGTACAGATATCAGCGGACCCCTCAAGTTTGGCTGTGAGACATT 420  
QY 778 GTGGAAGATACGAGATGAGATGAGTATTCGATTCCTCCAGAGAGGCTGCAACGTTAA 837  
DB 421 GTGGAAGATACGAGATGAGATGAGTATTCGATTCCTCCAGAGAGGCTGCAACGTTAA 480  
QY 838 GACAACTTTGCAAGTAAAGCGGACAGATCTATGTGACCATGCGCTGACAGATCTACAGAT 897  
DB 481 GACAACTTTGCAAGTAAAGCGGACAGATCTATGTGACCATGCGCTGACAGATCTACAGAT 540  
QY 898 GAGCTA 903  
DB 541 GAGCTA 546

Search completed: March 30, 2003, 18:47:39  
Job time : 261.709 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:39:57 ; Search time 59.2939 Seconds

(without alignments)  
5529.028 Million cell updates/sec

Title: US-10-082-502-18

Perfect score: 1069

Sequence: 1 gaattcgccgcaggggggt.....aaaaaatctcgccgcgc 1069

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapept 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/1na/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/1/1na/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/1/1na/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/1/1na/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/1/1na/PCTUS.COMB.seq:\*

6: /cgn2\_6/ptodata/1/1na/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	4.3	3581	2	US-08-738-349-1
2	45.6	4.3	7218	1	US-08-232-463-14
3	43.8	4.1	746	4	US-09-013-810-1
4	43.8	4.1	1780	4	US-09-202-5488-5
5	42.4	4.0	3238	4	US-08-123-934A-5
6	42.4	4.0	3238	5	PCT-US94-10080-5
7	42.2	3.9	3410	4	US-09-020-956-110
8	42.2	3.9	3410	4	US-09-030-607-110
9	42.2	3.9	3410	4	US-09-005-785-110
10	42.2	3.9	3410	4	US-09-352-616A-110
11	42.2	3.9	3410	4	US-09-602-877A-100
12	42.2	3.9	3410	4	US-09-232-149A-110
13	42.2	3.9	5852	1	US-07-867-106-2
14	42.2	3.9	6200	4	US-09-439-923-1
15	42.2	3.9	1114	4	US-09-152-060-41
16	42	3.9	1461	5	PCT-US95-04258-4
17	42	3.9	1864	4	US-09-149-476-130
18	42	3.9	289	1	US-08-341-568-3
19	41.6	3.9	289	2	US-08-911-020-3
20	41.6	3.9	1332	2	US-09-057-762-1
21	41.6	3.9	1332	3	US-08-326-119A-1
22	41.6	3.9	976	2	US-08-504-459-9
23	41.4	3.9	2205	3	US-08-888-077A-41
24	41.2	3.8	770	4	US-09-385-982-542
25	41	3.8	1248	2	US-08-897-340-5
26	41	3.8			
27	41	3.8			

28	41	3.8	1248	3	US-09-252-329-5	Sequence 5, Appl
29	41	3.8	1315	4	US-09-721-822A-10	Sequence 10, Appl
30	41	3.8	7452	3	US-08-592-500-1	Sequence 1, Appl
31	41	3.8	7452	3	US-08-195-006-1	Sequence 1, Appl
32	41	3.8	7452	5	PCT-US94-07644A-1	Sequence 1, Appl
33	40.8	3.8	458	1	US-08-524-757-1	Sequence 1, Appl
34	40.8	3.8	1502	4	US-09-206-903A-2	Sequence 1, Appl
35	40.8	3.8	1502	4	US-09-206-903A-11	Sequence 11, Appl
36	40.8	3.8	1502	4	US-09-202-122-2	Sequence 2, Appl
37	40.8	3.8	1502	4	US-09-202-122-1	Sequence 2, Appl
38	40.8	3.8	1502	4	US-09-206-935-2	Sequence 2, Appl
39	40.8	3.8	1502	4	US-09-206-935-3	Sequence 3, Appl
40	40.8	3.8	1502	4	US-09-206-936-2	Sequence 2, Appl
41	40.8	3.8	1502	4	US-09-206-936-3	Sequence 3, Appl
42	40.8	3.8	1582	3	US-08-545-196B-10	Sequence 10, Appl
43	40.8	3.8	1582	3	US-08-545-196B-12	Sequence 12, Appl
44	40.6	3.8	253	2	US-08-520-678A-25	Sequence 25, Appl
45	40.6	3.8	253	4	US-08-897-126-25	Sequence 25, Appl

## ALIGNMENTS

RESULT 1

US-08-738-349-1

Sequence 1, Application US/08738349

Patent No. 5869638

GENERAL INFORMATION:

APPLICANT: Takashita, Sunao

APPLICANT: Okazaki, Makoto

APPLICANT: Kawai, Shinji

APPLICANT: Tsujimura, Atsushi

APPLICANT: Amann, Egon

TITLE OF INVENTION: Bone-Related Cadherin-Like Protein and

NUMBER OF INVENTION: Process for Its Production

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/738,349

FILING DATE: 25-OCT-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/364,439

FILING DATE:

APPLICATION NUMBER: US 08/112,061

FILING DATE: 26-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Barker, M. P.

REGISTRATION NUMBER: 32,013

REFERENCE/DOCKET NUMBER: 02481.1323-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4400

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3581 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: Mus musculus  
STRAIN: osteoblastic cell line MC3T3E1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 284..2671  
US-08-738-349-1

Query Match 4.3%; Score 46; DB 2; Length 3581;  
Best Local Similarity 63.6%; Pred. No. 0.0031;  
Matches 70; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Y 960 AAGATGCGAGCATGCTTTATATAGCTTTTGTGAAATGAACTGAAAAACTCT 1019  
Db 3466 AAGAAGCAACCTTTAAATATAAATAAGATTCTTTTAAAAAATAAAAAA 3525

Y 1030 GAACCGAAGTAAAAAATTTCCGCGCGC 1069  
Db 3526 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 3575

RESULT 2  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367

GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEFFELINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232.463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935.313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-f18  
US-08-232-463-14

Query Match 4.3%; Score 45.6; DB 1; Length 7218;  
Best Local Similarity 5.0%; Pred. No. 0.0057;  
Matches 21; Conservative 219; Mismatches 178; Indels 0; Gaps 0;

Y 643 ATTGACCTTCTACCCAGCAAGTACGACGCTCGGAGCGGAGAGATCC 702  
Db 1481 AATTACCTATCTATCACTAGTAAAGATAGAAATTTGTAACATTTTATTTT 1422  
Y 703 AGTGAAGTACTTACGAGGATCCGAATGACTGAGATATGAGGAGCCCTCAGTT 762  
Db 1421 RRR 1362  
Y 763 GCGTGTGAGCATGTTGGAAGATGAGATGAGCTTATGCAATCTTCTCCAGAG 822  
Db 1361 RRR 1302  
Y 823 GGTGACAGCTTAAAGACAACTTTCAGTAAAGCAGATCTTGTGACATGCGCTG 882  
Db 1301 RRR 1242  
Y 883 CACAGATCTCAGATGAGTATGATCACTGAGCAGCAGCCTCACCAAGTATG 942  
Db 1241 RRR 1182  
Y 943 AACACCCCGAGGAGGAGATGAGCATGCTTTTATATAGCTTTTATGAAATG 1002  
Db 1181 RRR 1122  
Y 1003 AACTGAAAAAAGCTTTGAAACCGAAGTAAAAAATTTCC 1060  
Db 1121 RRR 1064

RESULT 3  
US-09-013-810-1  
Sequence 1, Application US/09013810  
Patent No. 6197551

GENERAL INFORMATION:  
APPLICANT: Busfield, Samantha J.  
TITLE OF INVENTION: No. 6197551a TANGO 80 Molecules and Uses Thereof  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/013.810  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragoras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MEI-010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 746 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 135..428





```

1  PRIOR APPLICATION DATA:
2  APPLICATION NUMBER:  US 08/123,934
3  FILING DATE:  17-SEP-1993
4  CLASSIFICATION:
5  ATTORNEY/AGENT INFORMATION:
6  NAME:  LAZAR, Steven R
7  REGISTRATION NUMBER:  32,618
8  REFERENCE/DOCKET NUMBER:  5203-PCT
9  TELECOMMUNICATION INFORMATION:
10 TELEPHONE:  (617) 498-8260
11 TELEFAX:  (617) 876-8651
12 INFORMATION FOR SEQ ID NO:  5:
13 SEQUENCE CHARACTERISTICS:
14     LENGTH:  3238 base pairs
15     TYPE:  nucleic acid
16     STRANDEDNESS:  single
17     TOPOLOGY:  linear
18 MOLECULE TYPE:  DNA (genomic)
19 IMMEDIATE SOURCE:
20 CLONE:  CFK1-10a
21 FEATURE:
22     NAME/KEY:  CDS
23     LOCATION:  474..2000
24
25 PCT-US94-10080-5

```

Query Match	4.08;	Score 42.4;	DB 5;	Length 3238;
Best Local Similarity	69.08;	Pred. No. 0.031;		
Matches	58;	Conservative	0;	Mismatches 0;
				Gaps 0

973 TGCCTTTTATTACGTTTTATGGAATGAACTGAAAACTCTGAAACCGAAAGTA 1032

Db 3106 TACCTTTTGTATTAGTACTTATTGTAATAAATACTGTTTTCAAGTCAAAAA 3165

QY	1033	AAAAAAAAAAAAAAAA	1056
----	------	------------------	------

Db 3166 AAAAAAAAAAAAAAAAAA 3189

RESULT 7  
DS-09-020-956-110

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
US-09-020-956-110

```

Query Match	3.98;	Score 42.2;	DB 4;	Length 3410;
Best Local Similarity	54.28;	Pred. No. 0	036;	
Matches	83;	Conservative	1;	Mismatches 69;
				Indels 0;
				Gaps 0.

904 TGAATCACTGGAGCAAGCCTACACCAACGTGATGGACACCCCCAGGAGGGAGA 963

Db 3236 TTATTAGCGGGTGAATATTTATACTGTAAGTGAGCAATCAGAGTATAATGTTATGG 3295

964 TGGCAGCATTTGCCCTTTTATATTACGTTTATTGGAATGAACTGAAAAAAGCTCTGAA 1023

DB 3296 TGGACAAATTAAGGCTTTCTTATATGTTTAAAAA 3355

```

QY 1024 CCGAAGTAAAAAAAAAAAAAAAAAAAAA 1056
      ||| ||||:|||||||
Db 3356 AAAAAAAAAAAAAAAAAAAAAAAAAA 3388

```

RESULT 8  
US-09-030-607-110  
; Sequence 110, Application US/09030607  
; Patent No. 6262245

Query Match	3.98	Score	42.2	DB	4	Length	3410
Best Local Similarity	54.28	Pred. No.	0.036				
Matches	83	Conservative	1	Mismatches	69	Indels	0
				Gaps			0

904 TGAATCACTGGAGCAAGCAGCCTACACCAACGTGATGGAACACACCCCCAGGAGGGAAGA 963

Db 3236 TTAATTAGCGGGGTGAATTTTATCTAAGTGAAGCAATGAGTAAATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAACTCTTGAAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAA 3355  
QY 1024 CCGAAGTAAAAAAAAAAAAAAAAAAAAA 1056  
Db 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 9

US-09-605-785-110  
Sequence 110, Application US/09605785  
Patent No. 6321716

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Reltter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darlick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aljun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C16  
CURRENT APPLICATION NUMBER: US/09/605,785  
CURRENT FILING DATE: 2000-06-27  
NUMBER OF SEQ ID NOS: 835  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-605-785-110

Query Match  
Best Local Similarity 3.9%; Score 42.2; DB 4; Length 3410;  
Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGAGCAAGCAGCTACACCAAGCGATGGAACACCCCGAGAGGGAAGA 963  
Db 3236 TTAATTAGCGGGGTGAATTTTATCTAAGTGAAGCAATGAGTAAATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAACTCTTGAAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAA 3355  
QY 1024 CCGAAGTAAAAAAAAAAAAAAAAAAAAA 1056  
Db 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 10

US-09-439-313-110  
Sequence 110, Application US/09439313  
Patent No. 6329505

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise

APPLICANT: Jiang Yuqi  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Reltter, Mark  
APPLICANT: Stolk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C9  
CURRENT APPLICATION NUMBER: US/09/439,313  
CURRENT FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-439-313-110

Query Match  
Best Local Similarity 3.9%; Score 42.2; DB 4; Length 3410;  
Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGAGCAAGCAGCTACACCAAGCGATGGAACACCCCGAGAGGGAAGA 963  
Db 3236 TTAATTAGCGGGGTGAATTTTATCTAAGTGAAGCAATGAGTAAATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAACTCTTGAAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAA 3355  
QY 1024 CCGAAGTAAAAAAAAAAAAAAAAAAAAA 1056  
Db 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 11

US-09-352-616A-110  
Sequence 110, Application US/09352616A  
Patent No. 6395278

## GENERAL INFORMATION:

APPLICANT: Dillon, Davin C.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang, Yuqi  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitcham, Jennifer Lynn  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.427C8  
CURRENT APPLICATION NUMBER: US/09/352,616A  
CURRENT FILING DATE: 1999-07-13  
NUMBER OF SEQ ID NOS: 472  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 110  
LENGTH: 3410  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-352-616A-110

Query Match  
Best Local Similarity 3.9%; Score 42.2; DB 4; Length 3410;  
Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGAGCAAGCAGCTACACCAAGCGATGGAACACCCCGAGAGGGAAGA 963  
Db 3236 TTAATTAGCGGGGTGAATTTTATCTAAGTGAAGCAATGAGTAAATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTTATATTAAGCTTTTATGAAATGAACTGAACTCTTGAAA 1023  
Db 3296 TGACAAATTTAAAGCTTCTTATATGTTTAAAAAAAAAAAAAAAAAAAAAA 3355  
QY 1024 CCGAAGTAAAAAAAAAAAAAAAAAAAAA 1056

Db 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 12

US-09-602-877A-100  
; Sequence 100, Application US/09602877A  
; Patent No. 6432707  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.446C5  
; CURRENT APPLICATION NUMBER: US/09/602.877A  
; CURRENT FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 100  
; LENGTH: 3410  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-602-877A-100

Query Match 3.9%; Score 42.2; DB 4; Length 3410;  
Best Local Similarity 54.2%; Pred. No. 0.036;  
Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGGAGCAGACGCTTACACCAACGATGGAACCCCGAGGGGGAAGA 963  
DB 3236 TTATTACGGGGGTGATATTTTATACGTGATGAGCAATCAGATATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTATATTACGTTTATGAAATGCAACTGAAAACTCTGAAA 1023  
DB 3296 TGACAAATTTAAGGCTTCTTATATGTTAAAAAAGAAAAAAAAAAAAAAAAA 3355  
QY 1024 CCGAAAGTAAAAAAGAAAAAAAAAAAAAAAAA 1056  
DB 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 13

US-09-232-149A-110  
; Sequence 110, Application US/09232149A  
; Patent No. 6465611  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer Lynn  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE  
; FILE REFERENCE: 210121.447C6  
; CURRENT APPLICATION NUMBER: US/09/232.149A  
; CURRENT FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 338  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 110  
; LENGTH: 3410  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-232-149A-110

Query Match 3.9%; Score 42.2; DB 4; Length 3410;  
Best Local Similarity 54.2%; Pred. No. 0.036;  
Matches 83; Conservative 1; Mismatches 69; Indels 0; Gaps 0;

QY 904 TGAATCACTGGAGCAGACGCTTACACCAACGATGGAACCCCGAGGGGGAAGA 963  
DB 3236 TTATTACGGGGGTGATATTTTATACGTGATGAGCAATCAGATATGTTATG 3295  
QY 964 TGGCAGCATTCCTTTATATTACGTTTATGAAATGCAACTGAAAACTCTGAAA 1023  
DB 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

Db 3296 TGACAAATTTAAGGCTTCTTATATGTTAAAAAAGAAAAAAAAAAAAAAAAA 3355  
QY 1024 CCGAAAGTAAAAAAGAAAAAAAAAAAAAAAAA 1056  
DB 3356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3388

## RESULT 14

US-07-867-106-2/C  
; Sequence 2, Application US/07867106  
; Patent No. 5389526  
; GENERAL INFORMATION:  
; APPLICANT: Slade, Martin B  
; APPLICANT: Chang, Andy C M  
; APPLICANT: Williams, Keith L  
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz MacKiewicz & No. 5389526rls  
; STREET: One Liberty Place 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/867.106  
; FILING DATE: 19920625  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PJ 7187  
; APPLICATION NUMBER: PCT/AU90/00530  
; FILING DATE: 02-NOV-1989

ATTORNEY/AGENT INFORMATION:  
; NAME: Feeney, Joanne Longo  
; REGISTRATION NUMBER: 35,134  
; REFERENCE/DOCKET NUMBER: RICE-0002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5852 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 2378..5038  
; NAME/KEY: CDS  
; LOCATION: 2378..5038  
; NAME/KEY: CDS  
; LOCATION: 2378..5038  
US-07-867-106-2

Query Match 3.9%; Score 42.2; DB 1; Length 5852;  
Best Local Similarity 70.9%; Pred. No. 0.047;  
Matches 56; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 978 TTATATTACGTTTATGGAATGACGTAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1037  
DB 5682 TTATATTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 5623  
QY 1038 AAAAAAAAAAAAAAAAAAAAAA 1056  
DB 5622 AAAAAAAAAAAAAAAAAAAAAA 5604

```

RESULT 15
US-09-439-923-1
; Sequence 1, Application US/09439923
; Patent No. 6426208
; GENERAL INFORMATION:
; APPLICANT: Emily D. Kakkis
; APPLICANT: Becky Tanamachi
; TITLE OF INVENTION: Recombinant Alpha-L-Iduronidase, Methods
; TITLE OF INVENTION: for Producing and Purifying the Same and Methods for
; FILE REFERENCE: 08000510500
; CURRENT APPLICATION NUMBER: US/09/439,923
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO. 1
; LENGTH: 6200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1558)...(3516)
US-09-439-923-1

```

```

Query Match          3.9%; Score 42.2; DB 4; Length 6200;
Best Local Similarity 67.8%; Pred. No. 0.049;
Matches 59; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

OY  970 CATGCCCTTTTATTACGTTTGTGGAATGAACTGAAAAAAGCTTGAAACCGAAA 1029
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  3608 CTTGCAATATATATTTTATATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 3667
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY  1030 GTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 1056
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  3668 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 3694
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 30, 2003, 20:59:50  
 Job time : 104.294 secs







PRIOR APPLICATION NUMBER: 60/059352	PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588	PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836	PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250	PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285	PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287	PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814	PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816	PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045	PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082	PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127	PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063329	PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550	PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561	PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704	PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733	PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735	PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738	PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755	PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248	PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809	PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186	PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846	PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364	PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453	PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511	PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770	PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212	PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278	PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334	PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694	PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320	PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612	PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086	PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092	PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/07791	PRIOR FILING DATE: 1998-02-09

PRIOR FILING DATE:	1998-03-12
PRIOR APPLICATION NUMBER:	60/078910
PRIOR FILING DATE:	1998-03-20
PRIOR APPLICATION NUMBER:	60/079294
PRIOR FILING DATE:	1998-03-25
PRIOR APPLICATION NUMBER:	60/079653
PRIOR FILING DATE:	1998-02-27
PRIOR APPLICATION NUMBER:	60/079728
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/080165
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/081203
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081229
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/081695
PRIOR FILING DATE:	1998-04-14
PRIOR APPLICATION NUMBER:	60/081817
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/081818
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/082999
PRIOR FILING DATE:	1998-04-24
PRIOR APPLICATION NUMBER:	60/083322
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION NUMBER:	60/083545
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/084600
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084627
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/084637
PRIOR FILING DATE:	1998-05-07
PRIOR APPLICATION NUMBER:	60/085149
PRIOR FILING DATE:	1998-05-12
PRIOR APPLICATION NUMBER:	60/085323
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085338
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085339
PRIOR FILING DATE:	1998-05-13
PRIOR APPLICATION NUMBER:	60/085579
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085697
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/085704
PRIOR FILING DATE:	1998-05-15
PRIOR APPLICATION NUMBER:	60/086414
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/086430
PRIOR FILING DATE:	1998-05-22
PRIOR APPLICATION NUMBER:	60/087106
PRIOR FILING DATE:	1998-05-28
PRIOR APPLICATION NUMBER:	60/088026
PRIOR FILING DATE:	1998-06-04
PRIOR APPLICATION NUMBER:	60/088730
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088741
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088810
PRIOR FILING DATE:	1998-06-10
PRIOR APPLICATION NUMBER:	60/088858
PRIOR FILING DATE:	19/98-06-11
PRIOR APPLICATION NUMBER:	60/089532
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089599
PRIOR FILING DATE:	1998-06-17
PRIOR APPLICATION NUMBER:	60/089907
PRIOR FILING DATE:	1998-06-18
PRIOR APPLICATION NUMBER:	60/089947
PRIOR FILING DATE:	1998-06-19
PRIOR APPLICATION NUMBER:	60/090349
PRIOR FILING DATE:	1998-06-23

PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

Query Match 64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1,7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

QY 12 GAGGGGGTCTGCTGCTGCGAGGCGCTCTAAAGCTGCTGCGCGAGAGTTTG 71  
DB 196 GACCGGGGCTCTGCTGCTGCGAGCCTTCCT-AAAGCTGCTGCTGCGCGAGAGTTTG 254  
QY 72 GAGGGGCGGCTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 129  
DB 255 GAGGGGCGGCTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 314  
QY 130 GCTTGAAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 188  
DB 315 GCTTGAAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 374  
QY 189 CTGGCGGAGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 248  
DB 375 GTGGCAGGCGGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 434  
QY 249 CTAGGGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 308  
DB 435 -TAGTTTCAGTTGAAGTTCAGTTGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 493  
QY 309 ACGGCGGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 368  
DB 494 ACGGCGGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 553  
QY 369 GGGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 428  
DB 554 GGGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 613  
QY 429 AGATTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 488  
DB 614 AGATTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 673  
QY 489 CGTGGACCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 548  
DB 674 CGTGGACCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 733  
QY 549 GTCAGTTGAGGAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 608  
DB 734 GTCAGTTGAGGAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 793  
QY 609 GGTGTGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 668  
DB 794 GGTGTGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 853  
QY 669 CTAGGAGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 788  
DB 854 CTAGGAGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 913  
QY 729 AATGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 788  
DB 914 AATGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 973  
QY 789 CGAGGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 848

DB 974 CGAGGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 1033  
QY 849 CAGTAAAGCGGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 908  
DB 1034 CAGTAAAGCGGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 1093  
QY 909 CACTGAGGAGGAGGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 968  
DB 1094 CACTGAGGAGGAGGAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 1148  
QY 969 GCATGCGCTT--TATTTACGTTTATGAAATGAGTTCGAAAAA 1013  
DB 1149 GCATGCGCTT--TATTTACGTTTATGAAATGAGTTCGAAAAA 1195

RESULT 2  
US-10-121-049-173  
Sequence 173, Application US/10121049  
Publication No. US2003002239A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Geriltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Smerwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C17  
CURRENT APPLICATION NUMBER: US/10/121, 049  
CURRENT FILING DATE: 2002-04-12  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-121-049-173

Query Match 64.6%; Score 691; DB 9; Length 1210;  
Best Local Similarity 84.4%; Pred. No. 1,7e-187;  
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

QY 12 GAGGGGGTCTGCTGCTGCGAGGCGCTCTAAAGCTGCTGCGCGAGAGTTTG 71  
DB 196 GACCGGGGCTCTGCTGCTGCGAGCCTTCCT-AAAGCTGCTGCTGCGCGAGAGTTTG 254  
QY 72 GAGGGGCGGCTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 129  
DB 255 GAGGGGCGGCTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 314  
QY 130 GCTTGAAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 188  
DB 315 GCTTGAAGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 374  
QY 189 CTGGCGGAGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 248  
DB 375 GTGGCAGGCGGCGGCGGCTGCGAGGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 434  
QY 249 CTAGGGTTCAGTTGAAGTTCAGTTGCGGCGGCTGCGAGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 308  
DB 435 -TAGTTTCAGTTGAAGTTCAGTTGCGGCGGCGGCGGCTGCGACGG--GACCGTCAGATCTCC 493

QY	309	ACGGCGGGCGGCGCGGAGCAGAGAGTGGACCGACCCGTTTAACTTAAATATGAAAGGCTG	368
Db	494	ACGGCGGGCGGCGCGGAGCAGAGAGTGGACCGACCCCATTTACGCTTAAATATGAAAGGCTG	553
QY	369	GGGTGTGCTAGCCCTACTTCTTTGGGGGGTCTGTGTGGAGACTGCGTGGGCTCGAAGAGCCA	428
QY	429	AGATCTACACGTGTGGACCTTGGACGGGCTCGGGTGAATGAATTAAGATGGGAAATGGCCG	488
Db	614	GGATCTCACACTGTGGACATGACAGGGCTCTGGGTGAATTAAGATGGGAAATGGCCCA	673
QY	489	CGTGGAGCCCCAAGAAAGCATTTCAGATGGGATCTTCCGAATCAATCCAGATGGCAGCA	548
Db	674	GGTGGAGCCCCAAGAAAGCATTTCAGATGGGATCTTCCGGATCAATCCAGATGGCAGCA	733
QY	549	GTCAGTTGTGGAGSTACCTTATGCGCGCTCAGAGGCCACCTCACAGAGTCTCTTGAGGA	608
Db	734	GTCAGTGTGTGGAGSTACCTTATGCGCGCTCAGAGGCCACCTCACAGAGTCTCTTGAGGA	793
QY	609	GGGTGTGACCGAATGAAGAGTACGGGGGAACAGATTCACCTTCACCCACCGCAGAA	668
Db	794	GAATGTGACCGAATGAAGAGTATGGGGGAACAGATTCCTTCACCCATTCGCAAGAA	853
QY	669	CTACGTACGCGTCTGTGAGCCGGAATGGAGATTCAGTGAATAAGACTTCACAGGATCCG	728
Db	854	CTACGTACGTGTGTGTGGGCGCGGAATGGAGATTCAGTGAATAAGACTTCAGAGCATCCG	913
QY	729	TAATGACTCAGATATTCAGCGCGCACCTCAAGATTTCGCGTGTGAGAGCATTTGTGGAAGATA	788
Db	914	TAATGACTCAGATATTTAGCGCGCACCTCAAGATTTCGCGTGTGAGAGCATTTGTGAGAGATA	973
QY	789	CGAGATGAGAGCTTATCGAATTTCTCCAGAGAGGCGTGAACAAGTTTAAAGACAACCTTG	848
Db	974	CGAGATGAGACTCTTATTAATTTCTTTCCGAGAGGCTGACAATGTTTAAAGACAACCTTGT	1033
QY	849	CAGTAAACGGGACAGATCTATGTGACCATGCGCTGCACAGATCTCAAGATGAGACTATGAT	908
Db	1034	CAGTAAACGGGACAGATCTTGTGACCATGCGCTGCACATATGCAATGATGAGACTATGAAAC	1093
QY	909	CACGTGAGAGCAAGACGCTTACACCMAACGATGATGAAACACCCCGAGAGGGGAAGATGGCA	968
Db	1094	CACGTGAGAGCAAGCCAC-----ACTGGCTGTGATGATCACCCCGAGAGGGGAATGTGTG	1148
QY	969	GCATTCGCTTTT--TATATACGTTTTTATGGAATTAACATGAGAAAAA	1013
Db	1149	GCATTCGCTTTATATATTTATGTTTTTACTGAATTAATGAGAAAAA	1195
RESULT 3			
US-10-123-904-173			
: Sequence 173, Application US/10123904			
: Publication No. US2003002238A1			
GENERAL INFORMATION:			
: APPLICANT: Baker, Kevin P.			
: APPLICANT: Beresini, Maureen			
: APPLICANT: DeForge, Laura			
: APPLICANT: Desnoyers, Luc			
: APPLICANT: Filvaroff, Ellen			
: APPLICANT: Geo, Wei-Qiang			
: APPLICANT: Geriltsen, Mary E.			
: APPLICANT: Goddard, Audrey			
: APPLICANT: Godowski, Paul J.			
: APPLICANT: Gurney, Austin L.			
: APPLICANT: Sherwood, Steven			
: APPLICANT: Smith, Victoria			
: APPLICANT: Stewart, Timothy A.			
: APPLICANT: Tumas, Daniel			
: APPLICANT: Watanabe, Colin K			
: APPLICANT: Wood, Ziemlin			
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			

	; TITLE OF INVENTION:	ACIDS ENCODING THE SAME
	; FILE REFERENCE:	P333ORLC54
	; CURRENT APPLICATION NUMBER:	US/10/123,904
	; CURRENT FILING DATE:	2002-04-16
	; Prior Application removed -	See File Wrapper or Palm
	; NUMBER OF SEQ ID NOS:	550
	; SEQ ID NO 173	
	; LENGTH:	1210
	; TYPE:	DNA
	; ORGANISM:	Homo Sapien
	US-10-123-904-173	
	Query Match	64.6%; Score 691; DB 9; Length 1210;
	Best Local Similarity	84.4%; Pred. No. 1.7e-187;
	Matches 850; Conservative	0; Mismatches 145; Indels 12; Gaps
Oy	12	GAGGGGGGCTCCGTGCCTCGAGAGGCCGCTCCTAAAGCTGCCTCGCGGCAGACTTGG 71
Dd	196	GACCGGGGCTCTGTGCTTTGGGACCCTTCCT-AAGCTGCTGTTCGGCAGAGTTTTG 254
Oy	72	GAGGGCGGGCTTAAGGCTACATTTGGTGGGGGCTCGCACGG--GACCCTGAGATTCTC 129
Dd	255	GAGGGCGGGTTTGGGGTGGGTCTGAATTGGGGCTCGACCGCAGCAAGCTGAGATCCC 314
Oy	130	GCTTAGTTCCTAGTTAAGTCCGGGAAGCTGGCCAGGCGGTC-ACTGCCACCTGAAC 188
Dd	315	GCTTAGGTACCGATTAGGCGTCAAGGGAGCTGGGTCAAGGCGTCCGCGGAGCACCCGTTGT 374
Oy	189	CTGGGGGAGCCGGAGCGCTGTGGAAAGCCGGGAGACGCCGCTTTTCCAGCCAGCTG 248
Dd	375	GTGGCAGCGCGGGAAGCGCTCTGGAGMAATCCGGCAGACCCCTGCTCCCTGCACGCAAGTGTG 434
Oy	249	CTAGGGTTTGGGACCCACAGAAAACAAAGTGAGAGTCCGGCTCTTTCCAGAGCCCTGGGCC 308
Dd	435	-TAGTTTCGGGAGCCACTGGGGCCCAAAGTGAAGATCCAGCGGTCTTCCAGCGCTTGGGCC 493
Oy	309	ACGGGGGGGCGCCGTGGGAGCAGAGGTGGAGCGACCCCTGTTACACTAAGAATGAAGGCTG 368
Dd	494	ACGGGGGGGCGCCCTGGGAGCAGAGGTGGAGCGACCCCATTTACGTTAAGATGAAGAGGCTG 553
Oy	369	GGGTGGGCTAGCCCTACTATTTTGGGGGCTCCTCGTGGGAAGTGCCTGGGCGTCGAAAGAGCCA 428
Dd	554	GGGTGGGCTAGCCCTGCTTCTGCGGGGCGCCGTGGGAAGCCGCTGGGCTCGGAGAGAGCCA 613
Oy	429	AGATTAACAACACTGTGAGAGCTTGCAGGGCTCTGTGTGATGAATTAGAGTGGAAAATTGCCCG 488
Dd	614	GGATTTCCACTGTGTGGAGATGCAAGGCGCTGTGTGATGAATAAGATGGAAATTGGCCCA 673
Oy	489	CGTGGACCCCCAAGAAGACCATTCAGATGGAGATCCTTCGCAATCAATCCAGATGGGAGCCA 548
Dd	674	GGTGGACCCCAAGAAGACCATTCAGATGGAGATCCTTCGCGATCAATCCAGATGGGAGCCA 733
Oy	549	GTCACTTGTGGAGGTACCTTATGCCCGCTCAGAGGCCCAACCTGCACAGAGTTGCTTGAGA 608
Dd	734	GTCACTGTGTGGAGGTACCTTATGCCCGCTCAGAGGCCCAACCTGCACAGAGTTGCTTGAGA 793
Oy	609	GSTGTGTACCGGAATGAAGAGGTACGGGGGAACAGATTGACCCCTTACCCACCGGAAGA 668
Dd	794	GATATGTACCGGAATGAAGAGGTATGGGGGAACAGATTGATCTTCCACCCACCGGAAGA 853
Oy	669	CTACGTACGCGTGTGAGCCCGGAATGGAGAAATCCAGTGAATAGACTTACAGAGGCATCG 728
Dd	854	CTACGTACGTGTAGTGGGCCCGGAATGGAGAAATCCAGTGAATAGACTTACAGAGGCATCG 913
Oy	729	AATTGACTCAGATATCAAGCGGCACCCCTCAAGTTTGCCTGTGAGAGCATTTGTGAAGAATA 788
Dd	914	AATGACCTCAGATATTAACCGGCACCCCTCAAGTTTGCCTGTGAGAGCATTTGTGAAGAATA 973
Oy	789	CGAGATATAGCTTATTCGAATTTCTTCCAGAGAGGCTGCACAACGTTTAAAGCAAACTTTG 848
Dd	974	CGAGATATACTCATTTGATTTCTTTTCCCGAGAGGCTGCACAATGTTTAAAGCAAACTTTG 1033
Oy	849	CAGTAAGCGGACAGATCTATGTGACCATGCGCTGCACAGATCTCAGATGAGCTATGAT 908



```

:
: CURRENT FILING DATE: 2002-06-19
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
:
: type: DNA
: ORGANISM: Homo Sapien
: OS-10-175-746-173

```

Query Match	Similarity	84.4%	Score 691;	DB 9;	Length 1210;
Best Local	Similarity	84.4%	Pred. No. 1,76-187;		
Matches	850;	Conservative	0;	Mismatches 145;	Indels 12; Gaps
QY	12	GAGGGGGCTCTCGCTCGAGGCGCTCTCTAAAGCTGCTGCTGCGGAGAGTTTG	71		
Db	196	GACCGGGCTCTGCTCTGCGAGACCTCTCTCT-AAAGCTGCTGTTGCGGAGAGTTTG	254		
QY	72	GAGGGGGGGCTTAGGCTCAGTTTGGGTGGGGGCTCGACGG--GACCTCAGATCTCC	129		
Db	255	GAGGGGGGGTGGGGGTGGGTGTGTGTGTGGGGCTCGACCGACAGCTGTGATCTCC	314		
QY	130	GCTTGTGAGCCCTAGTTAAAGTGGGGAGCTGGGGCAGGGGCTG-CTGGCCACCCGAGAC	188		
Db	315	GCTTAGGTACCAAGTTAGGCTCAGGGGAGCTGGGTCTAGCGGCTGCCGGGACACCCGCTGT	374		
QY	189	CTGGCGGAGCGCGAGCGCTCTGAGAAAGCCGGAGACAGCCCGCTTTTCCAGCCAGCTG	248		
Db	375	GTCGACGCGGGGAGAGCGCTGTGGAAATCCGGGAGAGCCGTGCTCCCGACCGACGATG	434		
QY	249	CTAGGGTTGGGACCCACAGAAACAAAGTGAAGTCCGGCTCTTCCAGAGCTGGGCC	308		
Db	435	-TAGTTTGGGGGACCACTGGGGCCAAAGTGAAGTCAAGCGCTCTTCCAGCCCTTGGGCC	493		
QY	309	ACGGGGGGCGCCGTGGGAGACAGAGGTGGAGCGACCCCTGTTACACTAAAGATGAAGGCTG	368		
Db	494	ACGGCGGGGGCCCTGGGAGCAGAGGTGGAGCGACACCCCATTTACGCTTAAGATGAAGGCTG	553		
QY	359	GGGTTGGCTAGCCCTACTTCTTGGGGGCTCTGCTGGAACTGCTGGGCTCGAAGAGCCA	428		
Db	554	GGGTTGGCTAGCCCTACTTCTTGGGGGCTCTGCTGGAACTGCTGGGCTCGAAGAGCCA	613		
QY	429	AGATTACACTGTGAGCTTGCAGGGCTCTGGTGGATGAATTAGAGTGGGAAATTGCCCC	488		
Db	614	GGATTCTCACACTGTGAGCATGCAAGGGCTCTGGTGGATGAATTAGAGTGGGAAATTGCCCC	673		
QY	489	CGTGACCCCAAGAGACCATTCAGATGGGATCCCTCCGAATTCATCCAGATGGGAGCCA	548		
Db	674	GCTGACCCCAAGAGACCATTCAGATGGGATCCCTCCGAATTCATCCAGATGGGAGCCA	733		
QY	549	GTCAGTTGTGAGGTACCTTATGCCGCTCAGAGGCCCCACCTCAGAGAGATTCTTGAGGA	608		
Db	734	GTCAGTTGTGAGGTACCTTATGCCGCTCAGAGGCCCCACCTCAGAGAGATTCTTGAGGA	793		
QY	609	GCTGTGTACCGAATGAAGAGTACGGGGACACAGATTGACCTTCTTACCCACCGCAGAA	668		
Db	794	GATATGTACCGGAGAGAGATATGGGGGACACAGATTGATCTTCCACCCATCGCAAGAA	853		
QY	669	CTAGCTAGCGGCTGAGACCCCGGAATGGAGAAATCCAGTGAATCAAGACTTCAGAGGATCCG	728		
Db	854	CTAGCTAGCTGAGTGGGCGCGGAATGGAGAAATCCAGTGAATCAAGACTTCAGAGGATCCG	913		
QY	729	AATTGACTCAGATATACAGCGGACCCCTCAAGTTTACGTGTGAGACATTTGGAAGATA	788		
Db	914	AATGACTCAGATATATACGGGACCCCTCAAGTTTACGTGTGAGACATTTGGAAGATA	973		
QY	789	CGAGAGTAGCTTATGAAATCTTCTCCAGAGAGGCTGCACAAGTTAAAGACAACTTGG	848		
Db	974	CGAGAGTAGCTTATGAAATCTTCTTCCGAGAGGCTGCACAATGTTAAAGACAACTTGG	1033		
QY	849	CAGTGAAGGAGACATTCATGTGACCATGTCCTGCACAGATTCACAGATGACTATGAAT	908		
Db	1034	CAGTGAAGGAGACATTCATGTGACCATGTCCTGCACATTCGATGATGACTATGAAT	1093		

Oy	CACGTGAGCAGCAGCGCTTACACCAAAAGTATGTGAACACCCCAAGAGGGAAATGCCA	968
Db		
Oy		
Db	CACGTGAGCAGCCAC-----ACGTGGTTATGTGATCACCCCAGAGGGAAATGCTG	1149
Oy	GCATTTCGCTT--TATATACGTTTTATGGAATGAATGAAAAA	1013
Db		
Oy	GCATTGCGCTTTTATATTTATGTTTTACTGAATTAATACGAAAAA	1195
Db		

```

RESULT 6
US-10-176-918-173
; Sequence 173, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresinski, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zenlin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C382
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US/10/176,918
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-918-173

Query Match      64.6%; Score 691; DB 9; Length 1210;
Best Local Similarity 84.4%; Pred. NO. 1.7e-187;
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps

OY 12 GAGGGGGTCTCTCGCTGCTCGAGGCGCTCTAAAGCTGCTGCTCGCGAGAGTTTG 71
Db 196 GACCGGGGCTCTGCTGCTGCTCGAGGCGCTCTAAAGCTGCTGCTCGCGAGAGTTTG 254
OY 72 GAGGGCGGGCTTAAAGGCTACGTTTGGTGGGGGCTCGCAGCG--GACCTTCAGATCTCC 129
Db 255 GAGGGCGGGGTTTGGGGTGGGCTGTGTGATTGGGCTCGCAGCGCAGCAGCTGAGTCCC 314
OY 130 GCTTAGGCGCTCAGTTAAGTGGCGGGAAGCTGGGCGAGCGGCTC-ACCTGCCACCTCGAAC 188
Db 315 GCTTAGGCTCAGTTAAGTGGCTCAGGGGAGCTGGGTGAGGGGGTGGCGGAGCACCCCGTGT 374
OY 189 CTGGCGGAGAGCGGAGCGCTCTGAGAAAGCCGGAGACGAGCCGTTTTCGCCAGCAGCTG 248
Db 375 GTGGAGAGCGGCGGAGAGCCCTCTGGAGAAATCCCGAGACAGCCCTGCTCCCTGCAAGCAGGTG 434
OY 249 CTAGGGTTTGGAGCCACAGAAAAACAAGTGAAGTCCGGCTGCTTTCAGAGAGCCCTGGGCC 308
Db 435 -TAGTTTTCGGGAGCCACTGGGGCCAAAGTGAGAGTCCAGCGGTCTTCCAGCGCTTGGGCC 493
OY 309 ACGGCGGGCGGCGCTGGGAGCAGAGGTGGAGGAGCCTCTGTTACACTTAAAGATGAAGGCTG 368
Db 494 ACGGCGGGCGGCGCTGGGAGCAGAGGTGGAGGAGCACCCTTACGCTTAAAGATGAAGGCTG 553

```

```

; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-173

Query Match      64.6%; Score 691; DB 9; Length 1210;
Best Local Similarity 84.4%; Pred. No. 1,7e-187;
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

Db 554 GGGTGGCTGGCCCTGCTTCTGGGGGCCCTGCTGGAAACCGCCCTGGGCTCGAGAGACCA 613
Qy 429 AGATCTACTGTGTGAGCTTGCAGGGCTGTGTGTGAATTAAGAGTGGAAATTTGCCCG 488
Db 614 GGAATCTCACTGTGTGAGCATGACAGGGCTGTGTGTGAATTAAGAGTGGAAATTTGCCCA 673
Qy 489 CGTGGACCCCAAGAACCTTTCAGATGGGATCTTCCGAATCAATCAATGGCAGCCA 548
Db 674 GGTGGACCCCAAGAACCTTTCAGATGGGATCTTCCGAATCAATCAATGGCAGCCA 733
Qy 549 GTGAGTGTGAGAGTACCTTATGCCCCCTGAGAGGCCACCTGACAGATGTTGCTGAGGA 608
Db 734 GTGAGTGTGAGAGTGTGCTTATGCCCCCTGAGAGGCCACCTGACAGATGTTGCTGAGGA 793
Qy 609 GGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCTTTTACCCACCGCAAGAA 668
Db 794 GATATGTGACCGGATGAAGAGTATGGGGAACAGATTGATCTTCCACCCATGGCAAGAA 853
Qy 669 CTAGCTACGGGTGTGAGCCCGGAATGGAATCAAGTGAATGACTTACAGGCGATCCG 728
Db 854 CTAGCTACGGGTGTGAGCCCGGAATGGAATCAAGTGAATGACTTACAGGCGATCCG 913
Qy 729 AATTGACTCAGATATCAGCGCACCCCTCAATGTTGGGTGAGAGGATTTGGAAGATA 788
Db 914 AATCGACTCAGATATTAAGCGCACCCCTCAATGTTGGGTGAGAGGATTTGGAAGATA 973
Qy 789 CGAGATGAGCTTATGAAATTTCTCCAGAGAGGCTGACACAGTTAAAGACAAACTTGG 848
Db 974 CGAGATGAGCTTATGAAATTTCTCCAGAGAGGCTGACACAGTTAAAGACAAACTTGG 1033
Qy 849 CAGTAAGCGGAGATCTATGACCATGCCCCGTCACAGATCTCAGAGTATGATGAAT 908
Db 1034 CAGTAAGCGGAGATCTTGTGACCATGCCCCGTCACAGATCTCAGAGTATGATGAAT 1093
Qy 909 CACTGAGCAAGCAGCTACACCAACGTCGATGGAACACCCCGAGAGGGAGATGCGA 968
Db 1094 CACTGAGCAAGCAGCTACACCAACGTCGATGGAACACCCCGAGAGGGAGATGCGT 1148
Qy 969 GCATTCGCCCTT--TATTTACGTTTATGGAATGACTGAAAAA 1013
Db 1149 GCATTCGCCCTTATTTATTTATGTTTACTGAAATTAAGTAAAAA 1195

RESULT 7
US-10-176-921-173
; Sequence 173, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176,921
; PRIOR FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550

```

```

; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-173

Query Match      64.6%; Score 691; DB 9; Length 1210;
Best Local Similarity 84.4%; Pred. No. 1,7e-187;
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

Qy 12 GAGGGGGTCTCCGCTGCTCCGAGAGCGCTCTTAAGCTGCTGCTGCGGAGACTTGG 71
Db 196 GACCGGGGCTCTTGTCTGCTGCGGAGCTCTCT--AAGCTCCTGTTGCGGCGAGACTTGG 254
Qy 72 GAGGGGCGGGCTTAAAGGTGCTTGGTGGGGGCTCGGACGG--GACCTCAGATCTCC 129
Db 255 GAGGGGCGGGTGTGGGGTGGCTGATTTGGGGCTCGGACCGACACCGCTGGAATGCC 314
Qy 130 GCTTAAAGTCTAGTAAAGTGCAGGAGCTGGGCGAGGGGCTC--ACTGGCCACCTGAAC 188
Db 315 GCTTAAAGTCTAGTAAAGTGCAGGAGCTGGGCGAGGGGCTGCGGCGACACCCGCTGT 374
Qy 189 CTGGCGGGAGCGGAGCGCTCTGTGAGAACCGGGACACGCCCTTTTCCAGACGCTG 248
Db 375 GTGGCAGGGGCGGAGCGCTCTGTGAGAACCGGGACACGCCCTCTCCTGACGCGAGTG 434
Qy 249 CTAGGCTTGGGAGCCACAGAAACAAAGTGAAGATCCGGCTGCTTCCAGAGCTTGGGG 308
Db 435 -TAGTTTCGGAGCCACCTGGGGGCGGAAAGTGAAGATCCAGCGCTTCCAGCGCTTGGGCC 493
Qy 309 ACGGCGGGCGCGCTGGGAGCAGAGTGAAGCGACCCCTGTTACACTTAAGATGAAGGCTG 368
Db 494 ACGGCGGGCGCGCTGGGAGCAGAGTGAAGCGACCCCTTACGCTTAAGATGAAGGCTG 553
Qy 369 GGGTGGCTAGCCCTACTTTTGGGGGCTCTGCTGGGGAATGCTGCGGCTCGAAGAGCCA 428
Db 554 GGGTGGCTAGCCCTACTTTTGGGGGCTCTGCTGGGGAATGCTGCGGCTCGAAGAGCCA 613
Qy 429 AGATCTACACTGTGTGAGCTTGTGAGGGCTGTGTGATTAAGATGAGTGGGAATTTGCCCG 488
Db 614 GGAATCTCACTGTGTGAGCATGAGGGCTGTGTGATTAAGATGAGTGGGAATTTGCCCA 673
Qy 489 CGTGGACCCCAAGAACCAATTCAGATGGATCTTCCGAATCAATCCAGATGGACCA 548
Db 674 GGTGGACCCCAAGAACCAATTCAGATGGATCTTCCGAATCAATCCAGATGGACCA 733
Qy 549 GTGAGTGTGAGAGTACCTTATGCCCCGTCAGAGGCCACCTCAGAGATCTCTGAGGA 608
Db 734 GTGAGTGTGAGAGTACCTTATGCCCCGTCAGAGGCCACCTCAGAGATCTCTGAGGA 793
Qy 609 GGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCTTCTACCCACCGCAAGAA 668
Db 794 GATATGTGACCGGATGAAGAGTATGGGGAACAGATTGATCTTCCACCCATGGCAAGAA 853
Qy 669 CTAGCTACCGCTGTGTGAGCGGGAATGGAATCCAGTGAATCAAGCTTACAGGGCATCG 728
Db 854 CTAGCTACCGCTGTGTGAGCGGGAATGGAATCCAGTGAATCAAGCTTACAGGGCATCG 913
Qy 729 AATTGACTCAGATATCAGCGCACCCCTCAATGTTGGGTGAGAGGATTTGGAAGATA 788
Db 914 AATCGACTCAGATATTAAGCGCACCCCTCAATGTTGGGTGAGAGGATTTGGAAGATA 973
Qy 789 CGAGATGAGCTTATGAAATTTCTCCAGAGAGCTGCAACGTTTAAAGAAACTTTG 848
Db 974 CGAGATGAGCTTATGAAATTTCTTCCGAGAGGCTGCAATGTTTAAAGAAACTTTG 1033
Qy 849 CAGTAAGCGGAGATCTATGACCATGCTGCAAGATCTCAGAGTATGATGAAT 908
Db 1034 CAGTAAGCGGAGATCTTGTGACCATGCTGCAAGATCTCAGAGTATGATGAAT 1093
Qy 909 CACTGAGCAAGCAGCTACACCAACGTCGATGGAACACCCCGAGAGGGAGATGCGA 968
Db 1094 CACTGAGCAAGCAGCTACACCAACGTCGATGGAACACCCCGAGAGGGAGATGCGT 1148

```







ORGANISM: HOMO Sapien  
US-10-140-474-173

Query Match	64.68;	Score 691;	DB 9;	Length 1210;
Best Local Similarity	84.48;	Pred. No. 1.7e-187;		
Matches 850; Conservative	0;	Mismatches 145;	Indels 12;	Gaps 6;

QY	12	GAGGGGGGCTCTCGGTGGCTCGAGAGGGGCTCTTAACCTGCGTGCCTGGCGGAGATTGG	71
Db	126	GACCGGGTCTCTGCTGTGTGGGAGCTCTTCCT--AACTGCTGTGTGGCGAGATTGG	254
QY	72	GAGGGGCGGCTTAAAGGCTACATTTCCGGTGGGGGGCTTCGACGG--GACCCTCAGATCTCC	129
Db	255	GAGGGCGCGGTTTTGGGGCTGGGTCTCTGATTTGGGGCTCGACACCGAGACGCTGGACTCC	314
QY	130	GCTTAAGTGGCTAGTTAAAGTCCGGGAAAGTGGGCTCAGAGCGGTC-AC TGCCACCTTGAAAC	188
Db	315	GCTTAAGTACCAATTTAGCGGTAGGGGAGCTGGGTCAAGGCGGTGCGGGGACACCCCGTGT	374
QY	189	CTGGCGGGAGCCGAGCGGCTTGTGAGAAAGCCGGGACAGCCCGCTTTTTCACACAGCTG	248
Db	375	GTCGACAGCGCGGAAGCGCTCTGGAGAAATCCCGGACAGCCCTCTCTCCCTGACCGAGTG	434
QY	249	CTAAGGCTTGGGACCCACAGAAACAAAGTAGAGTCCGGCTCTTTCAGAGGCTGGGCC	308
Db	435	-TAGTTTGGGAGGCTACTGTGGGGCCAAAGTAGAGTCCAGCGGTCCTTCAGCGCTTGGGCC	493
QY	309	ACGGCGGGCGGCGCTGTGGGAGCAGAGGTGGACGAGCCGCTTACCTAAAGTAGAAAGCTG	368
Db	494	ACGGCGGGGGGCGTGGGAGCGAGAGTGGACGACACCCATTACCTTAAAGTGAAGAGCTG	553
QY	369	GGGTTGGCTAGCCCTTACTTTTGGGGGTCTGTGGGAACTGCGTGGGCTCGAAGAGCCA	428
Db	554	GGGTTGGCTGGCCCTGCTTCTGGGGGGCCCTGTGGGAAACCGCTGGGCTCGGAGAGCCA	613
QY	429	AGATCTACACTGTGAGCTTGCAGAGGCGTGTGGTAATTAAGTGGGAAATTGGCCCG	488
Db	614	GGAATCTCACACTGTGAGCATATGCAAGGCTCTGTGTGATGAATGAATGGGAAATTGGCCCA	673
QY	489	CGTGACCCCAAGAACCAATTCAGATGGGATCTTCCGAAATCAATCAGATGGGAGCCA	548
Db	674	GGTGGACCCCAAGAAAGACCAATTCAGATGGGATCTTCCGAAATCAATCAGATGGGAGCCA	733
QY	549	GTCAGTTGTGGAGGTACCTTAATGCCGCTCAGAGGGCCACCTCACAGATTTGCTTGAAGA	608
Db	734	GTCAGTGTGGAGGTGCTTAATGCCGCTCAGAGGGCCACCTCACAGATTTGCTTGAAGA	793
QY	609	GGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCTTGCATCCACCGCAAGAA	668
Db	794	GATATGTGACCGGATGAAGAGTATGGGGAACAGATTGATCTTCCACCCATTCGCAAGA	853
QY	669	CTACGTACGGGTGTGAGCCGGGAATGGAAATTCAGTAATCAGACTTACAGGGCATCG	728
Db	854	CTACGTACGTGTGTGTGGCCGGGAATGGAATTCAGTAATCAGACTTACAGGGCATCG	913
QY	729	AATTGACTCAGATATCAGCGGCAACCTCAAGTTTGGGTGTGAGACATTTGTGAAGAAAT	788
Db	914	AATTGACTCAGATATTAGGGGCAACCTCAAGTTTGGGTGTGAGACATTTGTGAAGAAAT	973
QY	789	CGAGGATGAGCTTATGCAATTTCTTCCAGAGAGGCTGCAACGTTAAAGACAAACTTTG	848
Db	974	CGAGGATGAATCTATGAATTTCTTTTCCGAGAGGCTGCAATGTTAAAGACAAACTTTG	1033
QY	849	CAGTAAGCGGACAGATCTATGTGACCATGCGTGGACAGATATCAGATGAGCTTGAAT	908
Db	1034	CAGTAAGCGGACAGATCTTTGTBACATTCCTTGCACATATGCAATGATGAGCTTGAAC	1093
QY	909	CAGTGAGCAGACGAGCTTACACCAAACGTATGGAACACCCCGAGAGGGGAAGATGGCA	968
Db	1094	CAGTGAGCAGACGCCAC-----ACTGGCTGATGTGATCAACCCCGAGAGGGGAATGTG	1148
QY	969	GCAATGCGCTT--TATATTACGTTTTTATGGAATGAATCACTGAAAAA	1013

Db 1149 GCAATGCCCTTTTATATATTATGTTTTTACTGAATTACTGAAAAA 1195

**RESULT 10**

US-10-142-431-173

; Sequence 173, Application US/10142431

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

APPLICANT: Baker, Ke

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

**APPLICANT:** Gao, Wei-Qiang

APPLICANT: Gottard, Mary E

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy

APPLICANT: **Thomas, Daniel**

APPLICANT: WOOD W111327

APPLICANT: MOOU, WILLIAM

SECRET  
; TITLE OF INVENTION: SECRETED

**TITLE OF INVENTION: ACIDS E**

FILE REFERENCE: P3330R1C251

CURRENT APPLICATION NUMBER:

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

NUMBER OF SEQ ID NOS: 550

NUMBER OF SEQ ID NOS: 250  
; SEQ ID NO 173

LENGTH: 1210

; TYPE: DNA

ORGANISM: Homo Sapien

OS-10-142-431-173

Query Match 61 69

Query match	64.6%
Best local similarity	84.4%

Query Match	Score	DB %	Length
Best Local Similarity	84.6%	69.1	1210
Matches	850	Conservative	0
		Mismatches	145
		Indels	12
		Gaps	6

OY	489	CGTGGACCCCAAGAAAGACATTCAGATGGATCTCCGAAATCAATCCAGATGGCGACA	548
OY	489	CGTGGACCCCAAGAAAGACATTCAGATGGATCTCCGAAATCAATCCAGATGGCGACA	548
Db	674	GGTGGACCCCAAGAAAGACATTCAGATGGATCTTCCTGGATTAATCCAGATGGCGACA	733
OY	549	GTCAGTTGTGGAGGTACCTTATATGCCCCCTCAGAGGCCACCTCAGACAGATTGCTTAGNA	608
Db	734	GTCAGTTGTGGAGGTACCTTATATGCCCCCTCAGAGGCCACCTCAGAGAGTGTGGAGGA	793
OY	609	GCTGTGTACCCGATGAAGAGATACGGGGAAACAGATTGACCTTCTACCCACCGCAAGA	668
Db	794	GATATGTGACCCGATGAAGAGATATGGGGAAACAGATTGATTCCTCCACCATCGCAAGA	853
OY	669	CTTAGTACGCGTCTGTGAGCCGGAAATGAGAAATCCAGTGAAGTACAGATTACAGGGCATCCG	728
Db	854	CTACGTAAGTGTACTGTGGGCGCGGAATGAGAAATCAGTAAGTACAGTACAGAGCATCCG	913
OY	729	AATGTACTCAGATATACAGCGGACCCCTCAAGTTTGGCTGTGAGACGATTTGTGAGAAATA	788
Db	914	AATGTACTCAGATATATAGCGGACCCCTCAAGTTTGGCTGTGAGACGATTTGTGAGAAATA	973
OY	789	CGAGATGAGGCTTATTCGATTTCTTCACAGAGGGCTGACAAGTTTAAAGACAACTTGG	848
Db	974	CGAGATGAGGCTTATTCGATTTCTTCACAGAGGGCTGACAAGTTTAAAGACAACTTGG	1033
OY	849	CAGTAAAGCGGACAGATCTATGTGACCATGCGCTGCACAGATCTCAGATGAGCTATGAAT	908
Db	1034	CAGTAAAGCGGACAGATCTATGTGACCATGCGCTGCACAGATCTCAGATGAGCTATGAAT	1099
OY	909	CACGTGAGCAAGCAGCCTACACCAACGATGATGAACACCCCCAGAGGGGAAAGATGCA	968
Db	1094	CACGTGAGCAGCACCACCCACCCCTCAAGTTTATGATGATCACCACAGAGGGGAAAGATG	1144
OY	969	GCAATGCCCTTTTATATATGTTTATGGAATGAATGAAGAAATAA	1013
Db	1149	GCAATGCCCTTTTATATATGTTTATGGAATGAATGAAGAAATAA	1195
RESULT 11			
US-10-143-114-173			
: Sequence 173, Application US/10143114			
: Publication No. US20030036180A1			
: GENERAL INFORMATION:			
: APPLICANT: Baker, Kevin P.			
: APPLICANT: Beresini, Maureen			
: APPLICANT: Deforge, Laura			
: APPLICANT: Desnoyers, Luc			
: APPLICANT: Filvaroff, Ellen			
: APPLICANT: Gao, Wei-Qiang			
: APPLICANT: Geriltsen, Mary E.			
: APPLICANT: Goddard, Audrey			
: APPLICANT: Godowski, Paul J.			
: APPLICANT: Gurney, Austin L.			
: APPLICANT: Sherwood, Steven			
: APPLICANT: Smith, Victoria			
: APPLICANT: Stewart, Timothy A.			
: APPLICANT: Tumas, Daniel			
: APPLICANT: Watanabe, Colin K			
: APPLICANT: Wood, William			
: APPLICANT: Zhang, Zemin			
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
: FILE REFERENCE: P330301C211			
: CURRENT APPLICATION NUMBER: US/10/143,114			
: CURRENT FILING DATE: 2002-05-09			
: Prior Application removed - See Palm or File Wrapper			
: NUMBER OF SEQ ID NOS: 550			
: SEQ ID NO 173			
: LENGTH: 1210			
: TYPE: DNA			
: ORGANISM: Homo Sapien			
: US-10-143-114-173			

Query Match	64.68;	Score 691;	DB 9;	Length 1210;
Best Local Similarity	84.48;	Pred. No. 1.7e-187;		
Matches 850;	Conservative 0;	Mismatches 145;	Indels 12;	Gaps 6;

QY	12	GAGGGGGSTCTCGCTCCCTCGAGAGGGCGCTCAAGCTGACCTGGGCGGGAGAGTTTG	71
Dp	196	GACCGGGGCTCTTCTGCTTCTTGGAGGCTCTCTCT-AGCTGCTCTGTTCCGGCAGAGTTTG	254
QY	72	GAGGGGGGCGCTTAAGGCTCAGTTTCGTGGGGGGCTCGCACGG--GACCTTCAGATCTCC	129
Dp	255	GAGGGGGCGGTTTGGGGTCCGTGTCTGATTTGGGGGCTCGCACCGCACAGCGTGAGTCC	314
QY	130	GCTTAGTGGCTTAGTTAAGTCGGGAAAGCTGGGCGACGGCGTTC-AC TGGCCACCTTGAAAC	188
Dp	315	GCTTAGGTAGCAGTTAACGCTCAGGGGAGGCTGGGTCCAGGGGTGCGCGGGACACCCCTGT	374
QY	189	CTGCGCGGAGCCGGAGGCGCTCTGGAGAAAGCGGGAGACGCCGTTTTCCTCCAGCCAGCTG	248
Dp	375	GTCGACAGCGCGGCAAGCGCTCTGGAGAAATCCGGACAGCGCTGTCTCCCTGCAGCCAGGTG	434
QY	249	CTAAGTTGGGACCCACACAGAAAACAAAGTAGAGTCCGGCTCTTTCAGAGCCTGGGCC	308
Dp	435	-TAGTTTCGGGAGCCACTGGGGGCCAAAGTAGAGTCCACGGCTCTTCCAGCCTTGGGCC	493
QY	309	ACGGCGGGCGCGGGAGCAGAGGTGGAGCGACCTGTATACATMAAGATGAAGGCTG	368
Dp	494	ACGGCGGGCGCGCGGGAGCAGAGGTGGAGCGACCCCATATACGTCAAGATGAAGGCTG	553
QY	369	GGGTGGCTAGCCCTACTTTTGGGGTCCGTGCTGGAACTGCCTGGGCTCGAAGAGCCA	428
Dp	554	GGGTGGCTAGCCCTGCTTGTGGGGCCCTGCTGGGAACGGCTGGGCTCGSAGSAGCCA	613
QY	429	AGATCTCACTGTGGAGCTTCAGAGGCTCTGGTGATTAATTAAGTGGGAAATTTGCCG	488
Dp	614	GGAATCTCACTGTGGAGCATTGCAAGGCGCTGTGGTGTGAATAGATGGGAAATTTGCCA	673
QY	489	CGTGGACCCCAAGAAAGCCATTAGATGGGATGCTTCGAAATCATCCAGATGGCAGCCA	548
Dp	674	GGTGGACCCCAAGAAAGCCATTAGATGGGATGCTTCGGAATCAATTCAGATGGCAGCCA	733
QY	549	GTCAGTTGTGGAGGTACCTTATGCCGCTCAGAGGCCACCTCAGAGATGTTCTTGAGGA	608
Dp	734	GTCAGTTGTGGAGGTAGCTTATGCCGCTCAGAGGCCACCTCAGAGATGCTTGAGGA	793
QY	609	CGTGTGTACCCGAATGAAGAGTACGGGGGAACAGATTGACCTTGTACCCACCGCAACA	668
Dp	794	GATATGTGACCCGGAATGAAGAGTATGGGGAAACAGATTGATCTTCCACCCATCGCAAAA	853
QY	669	CTACGTACCGGTGTGTGAGCCGGAATGAGAAATCCAGTAACATAGACTTACAGGCGATCG	728
Dp	854	CTACGTACGTGTAGTGGGCGCGGAATGGAATTCAGTAACCTGACACTCAAGGATTCGG	913
QY	729	AATTGACTCAGATATCAGGGGCAACCTCAAGTTTGGGTGTGAGACATTTGTGAAGAATA	788
Dp	914	AATCGACTCAGATATTAGCGGCACCTCAAGTTTGGGTGTGAGACATTTGTGAGGAATA	973
QY	789	CGAGGATGACCTTATGGAATTTCTTCAGAGAGGCTGCAACGTTAAAGCAAACTTTG	848
Dp	974	CGAGGATGAACTATTTGAATTTCTTTCAGAGAGGCTGCAATGTTAAAGCAAACTTTG	1033
QY	849	CAGTAAACGAGACAGATCTATGTGTACCATCCCTGTGCACAGATCTCAGATGAGCTATGAT	908
Dp	1034	CAGTAAACGAGAAAGATCTTGTGTACCATCCCTGTGCACATGTGCATGATGAGCTATGAC	1093
QY	909	CAC TGGAGCAGACAGCCTACACCAAACTGTATGGAACACCCCGAGAGGGGAGATGSCA	968
Dp	1094	CAC TGGAGCAGCACCAC-----ACTGGCTGTATGTGATCAACCCCGAGAGGGGAAATGTG	1148
QY	969	GCATGCGCTT--TATATTACGTTTTTTATGAGAAATGAACGTGAAAAA	1013
Dp	1149	GCAATGCGCTTATATATTATGTTTTTACTGAAATTTAATCTGAAAAA	1195



```
OY 12 GAGGGGGCTCTGCTGCTCGAGGCGCTCTCTAAAGCTGCTCGCGGAGAGTTTG 71
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db GACCGGGGGCTCTGCTCGAGGCGCTCTCT - AAGCTGCTGTTCCGCGAGAGTTTG 254
OY 72 GAGGGGGGGCTTAAAGTCAAGTTGGTGGGGGGCTCGAGG--GACCTCAGATCTCC 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 GAGGGGGGGGGTGGGGTGGTCTGATTGGGGCTCGACCGACGACGCTGGAGTCC 314
OY 130 GCTTAAAGTCCCTAATTAAAGTGGCGGAAGCTGGCCAGCGGCTC-ACGCGCACCTGAAC 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 GCTTAAAGTCAAGTTAGGCTCAGGGGAGTGGGTCAAGGGGCTGGCGGAGACCCCTGT 374
OY 189 CTGGCGGAGCCCGGAGCGCTCTGAGAAAGCGGAGACGCCCTTTTCCAGCAGCTG 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 GTGGGAGCGCGGAGCGCTCTGAGAAATCCCGGACAGCCCTGCTCCGACAGCAGGTG 434
OY 249 CTAGGGTTGGGACCAAGAAACAAAGTGAAGTCCGGCTCTTTCCAGACCTCGGCC 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 -TAGTTTGGGAGCCACTGGGGCCAAAGTGAAGTCAAGCGGTCTTCCAGCGCTTGGGCC 493
OY 309 ACGCGCGGCGCGCTGGAGAGAGTGGAGCGACCCCTGTTACACTTAAAGTGAAGCTG 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 ACGCGGGGGGGCTTGGAGAGAGTGGAGCGACCCCAATTAGCTTAAAGTGAAGCTG 553
OY 369 GGGTGGCTAGCCCTACTTTTGGGGGCTCTGCGGAACCTGCTGGGCTCGAAGAGCCA 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 554 GGGTGGCTGGCGCTGCTTCTGGGGGCGCTGCGGAACCGCTGGGCTCGAGAGAGCCA 613
OY 429 AGATCTACAGTGGAGCTTGGAGGCTCTGCTGATGATGAATGAATGGGAATTTGCCG 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 GATCTTCACTGTGAGCATGAGAGGCTCTGCTGATGATGAATGAATGGGAATTTGCCA 673
OY 489 CGTGAGACCCCAAGAGACCATTCAGATGGATCTTCCGAATCAATCCAGATGGAGCCA 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 GGTGAGACCCCAAGAGACCATTCAGATGGATCTTCCGAATCAATCCAGATGGAGCCA 733
OY 549 GTCAATTTGGAGGTAAGTCTATGCGCGCTCAGAGGCCCACTCAGAGATTGCTTGAGGA 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 GTCAATTTGGAGGTAAGTCTATGCGCGCTCAGAGGCCCACTCAGAGATTGCTTGAGGA 793
OY 609 GGTGTGACCAAGATGAGAGAGTACGAGGGAACAGATTGACCTTTACCCACCGAGAA 668
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 GATATGTACCGGATGAGAGAGTATGGGGAACAGATTGATCTTCCACCCATCGCAAGAA 853
OY 669 CTACGTACCGCTCTGAGACCGGGAATGGAATCCAGTGAATCACTTACAGAGGCAATCG 728
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 CTACGTACGTGATGAGGCGCGGAATGGAATCCAGTGAATCACTTACAGAGGCAATCG 913
OY 729 AATTACACAGATTCAGAGCGCAACCTCAAGTTTCCGTGTGAGACATTTGGAAGATA 788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 AATTACACAGATTCAGAGCGCAACCTCAAGTTTCCGTGTGAGACATTTGGAAGATA 973
OY 789 CGAGATGAGCTTATCAATCTTCTCCAGAGAGCTGACAAAGTTTAAAGCAAACTTGG 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 974 CGAGATGAGCTTATCAATCTTCTCCAGAGAGCTGACAAAGTTTAAAGCAAACTTGG 1033
OY 849 CAGTAAAGCGAGATCTATGTACATGCTCTGACAGATCTCAAGTGAATGAT 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 CAGTAAAGCGAGATCTTGTGACATGCTCTGACATATGCGATGATGACATGTAAC 1093
OY 909 CACTGAGACCAAGACCCCTACACAAAGTGAATGGAACACCCCGAGAGGGAAGTGGCA 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 CACTGAGACCAAGACCCCTACACAAAGTGAATGGAACACCCCGAGAGGGAAGTGGTG 1148
OY 969 GCATTGCTTTT-TATATTACGTTTATGGAATGAGATGAAAAA 1013
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1149 GCATTGCTTTTATATTATGTTTACTGAAATTAAGTAAAAA 1195
```

RESULT 14  
US-10-123-262-173  
: Sequence 173, Application US/10123262

```
Publication No. US20030049816A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zhenlin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C38
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See File Wrapper or Palm
SEQUENCE ID NO: 173
LENGTH: 1210
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-262-173

Query Match 64.6%; Score 691; DB 9; Length 1210;
Best Local Similarity 84.4%; Pred. No. 1.7e-187;
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;

OY 12 GAGGGGGCTCTGCTGCTCGAGGCGCTCTCTAAAGCTGCTCGCGGAGAGTTTG 71
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 GACCGGGGGCTCTGCTCGAGGCGCTCTCT - AAGCTGCTGTTCCGCGAGAGTTTG 254
OY 72 GAGGGGGGGCTTAAAGTCAAGTTGGTGGGGGGCTCGAGG--GACCTCAGATCTCC 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 GAGGGGGGGGGTGGGGTGGTCTGATTGGGGCTCGACCGACGACGCTGGAGTCC 314
OY 130 GCTTAAAGTCCCTAATTAAAGTGGCGGAAGCTGGCCAGCGGCTC-ACGCGCACCTGAAC 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 GCTTAAAGTCAAGTTAGGCTCAGGGGAGTGGGTCAAGGGGCTGGCGGAGACCCCTGT 374
OY 189 CTGGCGGAGCCCGGAGCGCTCTGAGAAAGCGGAGACGCCCTTTTCCAGCAGCTG 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 GTGGGAGCGCGGAGCGCTCTGAGAAATCCCGGACAGCCCTGCTCCGACAGCAGGTG 434
OY 249 CTAGGGTTGGGACCAAGAAACAAAGTGAAGTCCGGCTCTTTCCAGACCTCGGCC 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 -TAGTTTGGGAGCCACTGGGGCCAAAGTGAAGTCAAGCGGTCTTCCAGCGCTTGGGCC 493
OY 309 ACGCGCGGCGCGCTGGAGAGAGTGGAGCGACCCCTGTTACACTTAAAGTGAAGCTG 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 ACGCGGGGGGGCTTGGAGAGAGTGGAGCGACCCCAATTAGCTTAAAGTGAAGCTG 553
OY 369 GGGTGGCTAGCCCTACTTTTGGGGGCTCTGCGGAACCTGCTGGGCTCGAAGAGCCA 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 554 GGGTGGCTGGCGCTGCTTCTGGGGGCGCTGCGGAACCGCTGGGCTCGAGAGAGCCA 613
OY 429 AGATCTACAGTGGAGCTTGGAGGCTCTGCTGATGATGAATGAATGGGAATTTGCCG 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 GATCTTCACTGTGAGCATGAGAGGCTCTGCTGATGATGAATGAATGGGAATTTGCCA 673
OY 489 CGTGAGACCCCAAGAGACCATTCAGATGGATCTTCCGAATCAATCCAGATGGAGCCA 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 GGTGAGACCCCAAGAGACCATTCAGATGGATCTTCCGAATCAATCCAGATGGAGCCA 733
OY 549 GTCAATTTGGAGGTAAGTCTATGCGCGCTCAGAGGCCCACTCAGAGATTGCTTGAGGA 608
```

```

Db 734 GTGAGTGTGAGGAGTGTGCTTATGCCCCGCTGAGAGGCCACCTTACAGAGCTGCTGGAGGA 793
OY 609 GGTGTGTGACCGAATGAGGAGTACGGGAGGAGATTTGACCTTTTACCCACCCGCAAGAA 668
Db 794 GATATGTGACCGGATGAGGAGTATGGGAGAGATTTGATGCTTCCACCATGGCAAGAA 853
OY 669 CTACGTACCGGTGTGTGACCCGGAATGGAGATTCAGTGAATCTAGACTTACAGGGCATCCG 728
Db 854 CTACGTACCGGTGTGTGAGCGCGGAGATGAGATCCAGTGAATCCAGCTTACAGGCAATCCG 913
OY 729 AATTGACTGAGATATCAGCGGCGACCTCAATGTTGGCTGTGAGAGCATTTTGGAGAGATA 788
Db 914 AATCGACTGAGATATGAGCGGCGACCTCAAGTTGGCTGTGAGAGCATTTTGGAGAGATA 973
OY 789 CGAGATGAGACTTATGCAATTTCTTCCAGAGAGGCTGACCAAGCTTAAAGCAAACTTTTG 848
Db 974 CGAGATGAGACTTATGCAATTTCTTCCAGAGAGGCTGACCAATGTTAAAGCAAACTTTTG 1033
OY 849 CAGTAAGCGGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGATGAGCTATGAAAT 908
Db 1034 CAGTAAGCGGACAGATCTTGTGTGACCATGCCCTGCACATGTCGATGATGAGCTATGAA 1093
OY 909 CACTGAGGACGAGGCTTACCAAGCGTATGAGAGACACCCCGAGAGGGGAGAGATGGCA 968
Db 1094 CACTGAGGACGAGGCTTACCAAGCGTATGAGAGACACCCCGAGAGGGGAGAGATGGTG 1148
OY 969 GCATTCGCTTTT--TATATACGTTTATGGAATGAAATGAACTGAAAAA 1013
Db 1149 GCATTCGCTTTTATATATATATGTTTTTACGAAATTAAGTAAAAA 1195

```

```

RESULT 15
US-10-142-423-173
: Sequence 173, Application US/10142423
: Publication No. US20030049817A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: DeForge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Geriltsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Guiney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C249
: CURRENT APPLICATION NUMBER: US/10/142,423
: CURRENT FILING DATE: 2002-05-10
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-142-423-173

```

```

Query Match 64.6%; Score 691; DB 9; Length 1210;
Best Local Similarity 84.4%; Pred. No. 1,7e-187;
Matches 850; Conservative 0; Mismatches 145; Indels 12; Gaps 6;
OY 12 GAGGGGGTCTGCTGCTCGGAGGCGCTTAAAGCTGCTGCTGCGCGGAGAGTTTG 71
II IIIIIII IIII IIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

```

```

Db 196 GACCGGGGCTTCTGCTGTGGAGACCTCTCT--MAGCTGCTGTGTCGGCGAGAGTTTG 254
OY 72 GAGGGGGGGCTTAGGGTCAAGTTTCGTTGGGGGGCTGCGACGG--GACCTTCAGATCTCC 129
Db 255 GAGGGGGGGTTTGGGGTGGTGTGATTTGGGGCTTCGACCCGAGACAGCTGAGAGTCC 314
OY 130 GCTTAGGCTTGTATGATGCGGAGAACTGGCCGACGGCGTC-AC TGCCACCTGAA 188
Db 315 GCTTAGGCTTGTATGATGCGGAGAACTGGGGTGTGAGAGCGGTGCGCGGACACCCGCTG 374
OY 189 CTGGCGGAGACCGGAGCGCTCTGAGAGAGCGGGGACAGCCCGCTTTTCCACGACGCTG 248
Db 375 GTGGCAGCGCGGCGCTGTGAGAGATCCGAGATCCGAGCGCTCTCTCTCTCTCTCTCTCT 434
OY 249 CTAGGGTGGGACCCACAGAAAAAGTAGAGATCGCGCTGCTTCCAGAGCGCTGGGCG 308
Db 435 -TAGTTTGGGAGCGCTGTGGGGCGCAAGGTAGAGATCCAGCGGCTTCCAGCGCTTGGGCG 493
OY 309 ACGGCGGCGCGCTGGGAGAGAGGTGAGCGACCCCTGTTTACACTAAAGATGAAAGCTG 368
Db 494 ACGGCGGCGCGCGCTGGGAGAGAGGTGAGCGACCCCTTACGCTTAAAGATGAAAGCTG 553
OY 369 GGGTGGCTAGCCCTACTTTTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 428
Db 554 GGGTGGCTAGCCCTACTTTTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 613
OY 429 AGATCTACACTGTGAGACTTGTGAGGCTGTGAGGATGATGATGATGATGATGATGATG 488
Db 614 GATCTACACTGTGAGACTTGTGAGGCTGTGAGGATGATGATGATGATGATGATGATG 673
OY 489 GGTGAGCCCGAGAGACATTCAGATGGATGATGATGATGATGATGATGATGATGATG 548
Db 674 GGTGAGCCCGAGAGACATTCAGATGGATGATGATGATGATGATGATGATGATGATG 723
OY 549 GTGAGTGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 608
Db 734 GTGAGTGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793
OY 609 GGTGTGTGACCGAATGAGGAGTACGGGAGGAGATTTGACCTTTTACCCACCCGCAAGAA 668
Db 794 GATATGTGACCGGATGAGGAGTATGGGAGAGATTTGATGCTTCCACCATGGCAAGAA 853
OY 669 CTACGTACCGGTGTGTGACCCGGAATGGAGATCCAGTGAATCTAGACTTACAGGGCATCCG 728
Db 854 CTACGTACCGGTGTGTGAGCGCGGAGATGAGATCCAGTGAATCCAGTGAATCTAGACTT 913
OY 729 AATTGACTGAGATATCAGCGGCGACCTCAATGTTGGCTGTGAGAGCATTTTGGAGAGATA 788
Db 914 AATCGACTGAGATATGAGCGGCGACCTCAAGTTGGCTGTGAGAGCATTTTGGAGAGATA 973
OY 789 CGAGATGAGACTTATGCAATTTCTTCCAGAGAGGCTGACCAAGCTTAAAGCAAACTTTTG 848
Db 974 CGAGATGAGACTTATGCAATTTCTTCCAGAGAGGCTGACCAATGTTAAAGCAAACTTTTG 1033
OY 849 CAGTAAGCGGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGATGAGCTATGAAAT 908
Db 1034 CAGTAAGCGGACAGATCTTGTGTGACCATGCCCTGCACATGTCGATGATGAGCTATGAA 1093
OY 909 CACTGAGGACGAGGCTTACCAAGCGTATGAGAGACACCCCGAGAGGGGAGAGATGGCA 968
Db 1094 CACTGAGGACGAGGCTTACCAAGCGTATGAGAGACACCCCGAGAGGGGAGAGATGGTG 1148
OY 969 GCATTCGCTTTT--TATATACGTTTATGGAATGAAATGAACTGAAAAA 1013
Db 1149 GCATTCGCTTTTATATATATATGTTTTTACGAAATTAAGTAAAAA 1195

```

Search completed: March 30, 2003, 22:23:46  
Job time : 109.633 secs

1000

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2003, 18:36:03 ; Search time 1522.26 Seconds

(without alignments)  
11373.235 Million cell updates/sec

Title: US-10-082-502-18

Perfect score: 1069  
Sequence: 1 gaattcgacgaggggggtc.....aaaaaaattccgcgcgcgc 1069

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlinu:\*  
5: em\_estrov:\*  
6: em\_estropl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	889.4	83.2	1255	11	AK019927 Mus muscu
2	852	79.7	853	11	AK013014 Mus muscu
3	835.4	78.1	992	13	AK013014 Mus muscu
4	817.2	76.4	856	13	AK013014 Mus muscu
5	813.8	76.1	891	13	AK013014 Mus muscu
6	771	72.1	945	13	AK013014 Mus muscu

7	770	72.0	772	11	AK007914 Mus muscu
8	745.8	69.8	754	11	AK013358 Mus muscu
9	744.6	69.7	802	12	BE008664 MY1_00160
10	734	68.7	853	12	BF535071 602054331
11	721	67.4	808	13	BF408525 602964995
12	712.8	66.7	863	13	BF454899 603173447
13	708.8	66.3	843	13	BF412539 602990995
14	686	64.2	709	13	BF411532 602965126
15	654.4	61.2	894	14	BQ963382 AGENCOURT
16	641	60.0	663	13	BF108692 602895165
17	635.4	59.4	888	13	BF412230 602990252
18	634	59.3	1199	12	BE539304 602054778
19	627.8	58.7	724	14	BQ210416 U1-R-DY1-
20	619.6	58.0	867	13	BF973691 602844540
21	612.6	57.3	731	12	BF532558 602074611
22	612.2	57.3	1117	13	BM552907 AGENCOURT
23	603.4	56.4	875	13	BF163950 603049829
24	600.6	56.2	609	10	AM825757 US22A07.Y
25	599.8	56.1	604	10	BE308275 601097890
26	599.8	56.1	880	14	BQ421827 AGENCOURT
27	591.4	55.3	678	10	BE309953 601091903
28	591.4	55.3	791	12	BE704443 602868633
29	591.4	55.3	914	12	BE799495 601589289
30	590.8	55.3	594	10	AM908970 U40906.X
31	590.6	55.2	868	14	BQ421069 AGENCOURT
32	585.8	54.8	887	13	BF162759 603048045
33	585.6	54.8	600	10	AM413913 uc96c08.X
34	582.8	54.5	596	10	BE134206 uc917c09.X
35	582.2	54.5	783	13	BF550436 uc9192905
36	581.4	54.4	595	10	AM414039 uc96g03.X
37	576.4	53.9	742	12	BE824549 602728454
38	573.2	53.6	656	12	BF287107 EST451698
39	571	53.4	783	13	BF163784 603049729
40	568.4	53.2	570	10	AM826122 uc22a07.X
41	567.8	53.1	607	12	BE556857 U1-R-BG2-
42	567.4	53.1	697	12	BF344334 602014708
43	566.2	53.0	657	14	BQ211867 U1-R-DY1-
44	563.8	52.7	665	10	AM919569 EST350873
45	558	52.2	776	13	BF522523 603175203

## ALIGNMENTS

RESULT 1	AK019927	1255 bp	mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:5330432A10;transmembrane protein 4, full insert sequence.
LOCUS	AK019927		
DEFINITION	AK019927		
ACCESSION	AK019927		
VERSION	AK019927.1		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (Strain: C57BL/6J) adult male pituitary gland cDNA to mRNA, clone: lib: RIKEN full-length enriched mouse cDNA library clone: 5330432A10.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1		
TITLE	Carinci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS	2		
TITLE	Carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	2049374		
	11042159		







Query Match	Best Local Similarity	79.7%	Score 852;	DB 11;	Length 853;
Matches 852;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
146	AAGTCCGGGAAGCTGGGGGAGGGGGTCACTGCGCCACCTGAAACCTGGGGGGAGCGGGAGC	205	putative transmembrane protein 4"		
2	AAGTCCGGGAAGCTGGGGGAGGGGGTCACTGCGCCACCTGAAACCTGGGGGGAGCGGGAGC	61	/codon_start=1 /protein_id="BAB28597.1" /db_xref="GI:12850115" /db_xref="MGD:MG1:1928477" /translation="MKMGWLLALLGLVLTCTAARSRSDIHCCACRALVDELEIAR VDPKRTTQKMSFRINPDGSSQSVYEVPAKREAHITELIEVCIDMKRTGQIDPSTHR KNYRVRRNGESELDDGIRIDSDISGLTKFACESIVEEYEDLEIEFSREADNVK DKLCSKRTDLCIDNALHMSDEL"		
214	a	209	c	261	g
261		169	t		
BASE COUNT	214	a	209	c	261
ORIGIN					

Db	782	TACACCAAAAGCTGATGGAAACCCCGACGAGGAGGGAAGATGGACAGATTCCTTTATATT	841
Qy	986	ACGTTTTATGCG	997
Db	842	ACGTTTTATGCG	853
RESULT 3			
LOCUS	BI456021		
DEFINITION	BI456021	992 bp	mRNA
ACCESSION	BI456021		linear
VERSION	BI456021.1		EST 21-AUG-2001
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus.		
AUTHORS	NIH-MGC		
TITLE	http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgsapb-remail.nih.gov		
	Tissue Procurement: Lothar Henninghausen Ph.D., Robin Humphreys		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMNL at:		
	http://image.llnl.gov		
	Plate: LMNL1631	row: c	column: 17
	High quality sequence stop: 854.		
FEATURES			
SOURCE	location/Qualifiers		
	1..992		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:5249992"		
	/clone_id="NCI_CGAP_Mam5"		
	/tissue_type="tumor, gross tissue"		
	/dev_stage="7 months"		
	/lab_host="DH10B"		
	/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Henninghausen/Robin Humphreys, NIH"		
BASE COUNT	256 a	246 c	295 g
ORIGIN		194 t	1 others
Query Match	78.1%	Score 835.4;	DB 13; Length 992;
Best Local Similarity	97.4%	Pred. No. 2e-123;	
Matches 923; Conservative	0;	Mismatches 17;	Indels 8; Gaps 7;
Qy	103	GGGCTGCGACGAGGACCTCAGATCTCGGCTTAGGTGCTAGTTAAGTGGGGAAGCTGGG	162
Db	6	GCGATCGGACGAGGACCTCAGATCTCGGCTTAGGTGCTAGTTAAGTGGGGAAGCTGGG	65
Qy	163	CCAGGCGGTACTGCGCACCCTGAACCTGGCGGGAAGCCGAGCCCTGTGGAAACCGGG	222
Db	66	CCAGGCGGTACTGCGCACCCTGAACCTGGCGGGAAGCCGAGCCCTGTGGAAACCGGG	125
Qy	223	ACAGCCCGGTTTTCCAGCGACGCTGAGTGAAGGCTGGGACCCACAGAAACAAAGTGAAG	282
Db	126	ACAGCCCGGTTTTCCAGCGACGCTGAGTGAAGGCTGGGACCCACAGAAACAAAGTGAAG	185
Qy	283	TCCGGCTGCTTTCCAGAGACCTGGGCAACGCGGCGCGCGTGGAGACAGAGTGAAGCAG	342
Db	186	TCCGGCTGCTTTCCAGAGACCTGGGCAACGCGGCGCGCGCTGGAGACAGAGTGAAGCAG	245
Qy	343	CCGTGTAACATAAAGATGAAGGCTGGGGTGGCTAGCCCTACTATTTTGGGGGCTCTG	402

```

DB 246 CCTGTACCTTAAGTGAAGGCTGGGCTTGGCTAAGCCCTTCTTTGGGGTCTCTGCTG 305
OY 403 GGAAGTGGCTGGGCTGGAAGAGCCAGATCTACACTGTGTGAGCTTGAGGCTCTGTGTG 462
DB 306 GGAAGTGGCTGGGCTGGAAGAGCCAGATCTACACTGTGTGAGCTTGAGGCTCTGTGTG 365
OY 463 GATGAATTAAGATGGGAATTTCCCGCGTGTGAGCCCAAGAACCAATTCAGATGGGATCC 522
DB 366 GATGAATTAAGATGGGAATTTCCCGCGTGTGAGCCCAAGAACCAATTCAGATGGGATCC 425
OY 523 TTCCCAATCAATCCAGATGGCAGCCAGTCACTGTGTGAGCTTATGCGCGCTCAGAG 582
DB 426 TTCCCAATCAATCCAGATGGCAGCCAGTCACTGTGTGAGCTTATGCGCGCTCAGAG 485
OY 583 GCCCACTCAGAGATGGTGTGAGGAGGTGTGTGACCGAATGAGAGAGTACGGGGAAACAG 642
DB 486 GCCCACTCAGAGATGGTGTGAGGAGGTGTGTGACCGAATGAGAGAGTACGGGGAAACAG 545
OY 643 ATTGACCTTCTACCCACCGCAAGACTACGCTGCTGTACCGCGGAATGAGAGATCC 702
DB 546 ATTGACCTTCTACCCACCGCAAGACTACGCTGCTGTACCGCGGAATGAGAGATCC 605
OY 703 AGTGAAGTACACTTACAGAGGATCCGAATTCAGATATGAGGCGACCCCTCAAGTTT 762
DB 606 AGTGAAGTACACTTACAGAGGATCCGAATTCAGATATGAGGCGACCCCTCAAGTTT 665
OY 763 GCGTGTGAGAGATGTGTGAGAGATACGAGATGAGCTTATGAAATCTTCTCAGAGAG 822
DB 666 GCGTGTGAGAGATGTGTGAGAGATACGAGATGAGCTTATGAAATCTTCTCAGAGAG 725
OY 823 GCTGACAA--CGTTAAAGACAACTTTGCAAGTAAAGGACAGATCT--ATGTACACATGCC 879
DB 726 GCTGACAAAGCTTTAAAGACAACTTTGCAAGTAAAGGACAGATCT--ATGTACACATGCC 785
OY 880 CTGCACAGATCTCAGAGATGAGC--TATGAATCACT--GGAGCAGACGCGCTACACCAAGCT 937
DB 786 CTGCACAGATCTCAGAGATGAGC--TATGAATCACT--GGAGCAGACGCGCTACACCAAGCT 844
OY 938 GATGGAACACCCCGCAGAG--GGGGAAGATGAGCAGATGAGCTTATATATACCTTTTATG 996
DB 845 GATGGAACACCCCGCAGAGAGGAGGAGATGAGCAGATGAGCTTATATATACCTTTTATG 904
OY 997 GAAAGTACATGAAAAAACTCTTGAACCGGAAGTAAAAAATTTTATATATATATATATG 1044
DB 905 GATTGAACCTG--AAAAAATCTTTGAACCGGAAGTAAAAAATTTTATATATATATATG 951

RESULT 4
BI454180 856 bp mRNA linear EST 21-AUG-2001
LOCUS DEFINITION 603170602P1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5250001 5',
ACCESSION BI454180
VERSION BI454180.1 GI:15244836
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 856)
AUTHORS NIH-MGC http://hgsc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

Plate: LLM11631 row: d column: 02
High quality sequence stop: 826.
FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="5250001"
/clone_id="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site: 1; Salt:
Site: 2; Nci1; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 207 a 217 c 266 g 166 t
ORIGIN
Query Match 76.4%; Score 817.2; DB 13; Length 856;
Best Local Similarity 99.3%; Pred. No. 1.7e-120;
Matches 852; Conservative 0; Mismatches 3; Indels 3; Gaps 3;
OY 93 TTTCCGTTGGGGGCTGCGACGGGACCTCAGATCTCCGTTAGTGTGCTAGTTAAGTGG 152
DB 1 TTTCCGTTGGGGGCTGCGACGGGACCTCAGATCTCCGTTAGTGTGCTAGTTAAGTGG 60
OY 153 GGAAGTGGGGGCGGACCTGCGACCCGGAACCTTGGGTTGGGAGCCGAGCCGCTCTGG 212
DB 61 GGAAGTGGGGGCGGACCTGCGACCCGGAACCTTGGGTTGGGAGCCGAGCCGCTCTGG 120
OY 213 AGAAGCGGGAGACGCCCTTTTCCAGCAGCTGCTAGGTTGGGAGCCACAGAAAC 272
DB 121 AGAAGCGGGAGACGCCCTTTTCCAGCAGCTGCTAGGTTGGGAGCCACAGAAAC 180
OY 273 AAAGTACAGTCCGCTCTTTCAGAGCTTGGCCAGCGCGCGCGCTGGAGACAGAG 332
DB 181 AAAGTACAGTCCGCTCTTTCAGAGCTTGGCCAGCGCGCGCGCTGGAGACAGAG 239
OY 333 GTGGAGCGACCTGTTACCTAAAGATGAAGGCTGGGTTGGCTAGTCCCTATTGGG 392
DB 240 GTGGAGCGACCTGTTACCTAAAGATGAAGGCTGGGTTGGCTAGTCCCTATTGGG 299
OY 393 GGTCTGCTGGGAGTCCCTGGGCTGGAAGAGCCAAAGATCTACACTGTGAGCTTGACG 452
DB 300 GGTCTGCTGGGAGTCCCTGGGCTGGAAGAGCCAAAGATCTACACTGTGAGCTTGACG 359
OY 453 GGTCTGCTGATGAATTAAGTGGGAATTTCCCGCTGGAACCCCAAGAACCATTTCA 512
DB 360 GGTCTGCTGATGAATTAAGTGGGAATTTCCCGCTGGAACCCCAAGAACCATTTCA 419
OY 513 GATGGGATCCCTCCGAATATCCAGATGGGACGAGTGTGGAGAGTACCTTATGCG 572
DB 420 GATGGGATCCCTCCGAATATCCAGATGGGACGAGTGTGGAGAGTACCTTATGCG 479
OY 573 CCGCTCAGAGGCCCACTCAGAGAGTGTGAGAGGTGTGACCGAATGAAGAGTGA 632
DB 480 CCGCTCAGAGGCCCACTCAGAGAGTGTGAGAGGTGTGACCGAATGAAGAGTGA 539
OY 633 CGGGGAACAGATTGACCTTCTACCCACCGCAAGAACTACGTAACGGCTGTGACCGGAA 692
DB 540 CGGGGAACAGATTGACCTTCTACCCACCGCAAGAACTACGTAACGGCTGTGACCGGAA 599
OY 693 TGGAGATCCAGTGAAGTACCTACAGGCGATCCGAATTTGACTCAGATATACGGGAC 752
DB 600 TGGAGATCCAGTGAAGTACCTACAGGCGATCCGAATTTGACTCAGATATACGGGAC 659
OY 753 CCTCAAGTTGGGTGAGAGCATTTGTGAAGAAATACAGAGATGAGCTTATGAAATCTT 812
DB 660 CCTCAAGTTGGGTGAGAGCATTTGTGAAGAAATACAGAGATGAGCTTATGAAATCTT 719
OY 813 CTCAGAGAGGCTGACAACTTAAAGACAACTTTGCAAGTAAAGCGAGACGATCTATGTA 872

```

Db	720	CTCCGACAGAGCGTGACACGTTAAAGACAACCTTTGCAGTAAAGCGGACAGATCTATGTGA	779
Oy	873	CCATCCCTGCACAGATTCACGATGACCTATGATCTACTGGAGCAACGACCTACACCA	932
Db	780	CCATCCCTGTG-ACATATTCACGATGACCTATGATCTACTGGAGCAACGACCTACACC	837
Oy	933	AACGTGATGAACACCC 950	
Db	838	AACGTGATGAACACCC 855	
RESULT 5			
BI408381			
LOCUS	BI408381	891 bp	mRNA linear EST 14-NOV-2001
DEFINITION	602964013p1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119563 5',		
ACCESSION	BI408381		
VERSION	BI408381.1	GI:15169304	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 891)		
JOURNAL	NIH-MC http://mhc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-remail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima		
	Bonaldo, Ph.D.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/INL at:		
	http://image.lnl.gov		
	Plate: LLM11291 row: 1 column: 04		
	High quality sequence stop: 873.		
FEATURES			
source			
	1..891		
	/organism="Mus musculus"		
	/strain="CZECH II"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:5119563"		
	/clone_1lb="NCI_CGAP_Lu33"		
	/tissue_type="pooled lung tumors"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: lung; Vector: pRT30-Pac (Pharmacia) with a		
	modified polylinker; Site:1: NotI; Site:2: EcoRI; 1st		
	strand cDNA was prepared from mRNA obtained from pooled		
	lung tumors with a Not I - oligo(dT) primer [5		
	TCGTACCAATCTGAAGTGGAGCGCGCTCTGTTTCTTTTCTTTT 3'].		
	Double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Not I and cloned into the Not		
	I and Eco RI sites of the modified pRT30 vector. Library		
	went through one round of normalization, and was		
	constructed by Bento Soares and M. Fatima Bonaldo. "		
BASE COUNT	240 a	215 c	258 g
ORIGIN			178 t
Query Match	76.1%;	Score 813.8;	DB 13; Length 891;
Best Local Similarity	97.6%;	Pred. No. 5.6e-120;	
Matches 858:	Conservative 0;	Mismatches 17;	Indels 4;
		Gaps 3;	
Oy	174	CTGGCCACCTGAACTTGGCGGAGCCGCGCTGTGGAAGACCGGAGACCCCGTT	233
Db	14	CTGGCCACTGTGAACCTTGGCGGAGCCGCGCTGTGGAAGACCGGAGACCCCGTT	73
Oy	234	TTTCCACACGCTGCTGAGGTTGGGACCCACAGAAAACAAAGTAAAGTCCGGCTGT	293
Db	74	TTTTCCACACGCTGCTGAGGTTGGGACCCACAGAAAACAAAGTAAAGTCCGGCTGT	133

OY	294	TCCAGAGCCTGGGCGCACGGCGCGGCCGTGGAGACGAGATGAGACGACCCTTTAACT	353
Db	134	TCCAGAGCCTGGGCGCACGGCGCGGCCGTGGAGACGAGATGAGACGACCCTTTAACT	193
OY	354	AAAGATGAAGAAGCGTGGGATTGGCTAGCCCCCTACTTTTGGGGGTCCTCTGGAACTGCCTG	413
Db	194	AAAGATGAAGAAGCGTGGGATTGGCTAGCCCCCTACTTTTGGGGGTCCTCTGGAACTGCCTG	253
OY	414	GCGCTGGAAGGAGCCAAAGATCTACACTGTGTGAGCACTTGCAAGGCTCTGTGTGATGAATTAGA	473
Db	254	GCGCTGGAAGGAGCCAAAGATCTACACTGTGTGAGCACTTGCAAGGCTCTGTGTGATGAATTAGA	313
OY	474	GTGGGAAATTTGCCGCGTGGAGACCCCAAGAAGACCATTCAGATGGGGATCCCTTCGGAATCAA	533
Db	314	GTGGGAAATTTGCCGCGCGTGGAGACCCCAAGAAGACCATTCAGATGGGGATCCCTTCGGAATCAA	373
OY	534	TCCAGATGGCAGCCACGTCAGTGTGGAGGTACCTTAATGCCCGCTCAGAGAGCCACCTCAC	593
Db	374	TCCAGATGGCAGCCACGTCAGTGTGGAGGTACCTTAATGCCCGCTCAGAGAGCCACCTCAC	433
OY	594	AGAGTTGCTTGAAGAGGTGTGTGACCCGAATGAAGAGTACGGGGGAACAAGATTGACCTTC	653
Db	434	AGAGTTGCTTGAAGAGGTGTGTGACCCGAATGAAGAGTACGGGGGAACAAGATTGACCTTC	493
OY	654	TACCCACGCGAAGACTACGTACGGCGTGTGACCGCGGAATGGAGAATCCAGTGAATAGATA	713
Db	494	TACCCACGCGAAGACTACGTACGGCGTGTGACCGCGGAATGGAGAATCCAGTGAATAGATA	553
OY	714	CTTACAGGGGCATCCGAATTGACTCAGATATACAGCGGACACCTTCAAGTWTGCG-TGTGAGA	772
Db	554	CTTACAGGGGCATCCGAATTGACTCAGATATACAGCGGACACCTTCAAGTWTGCGTGTGAGA	613
OY	773	GCATTGTGGAAATAACGAGGATGAGCTATACGAATTCCTCTCCAGAGAGCGTCGCAACG	832
Db	614	GCATTGTGGAAATAACGAGGATGAGCTATACGAATTCCTCTCCAGAGAGCGTCGCAACG	673
OY	833	TTTAAAGACAAACTTTTGCAGTAGACGAGACAGATCTATGTGACCATGCCCCGACAGATCTC	892
Db	674	TTTAAAGACAAACTTTTGCAGTAGAGGAGACAGATCTATGTGACCATGCCCCGACAGATCTC	733
OY	893	ACGATGAGCTATGAATACACTGTGAGACAGACGCTTACACCAACGATGGAACACCCCACA	952
Db	734	ACGATGAGCTATGAATACACTGTGAGACAGACGCTTACAC -AACGTGATGGAACACCCCACA	792
OY	953	GGAGGGG--AAGATGAGCAGCATTTGCTTTATATACGTTTTATGGAATAAGACTGAAA	1010
Db	793	GGAGGGGAGCAGCATTTGCTTTATATACGTTTTATGGAATAAGACTGAAA	852
OY	1011	AAAACTCTTGAACCGAATAATAAAAAAAAAAAAAAAAAAAAA	1049
Db	853	AAAACTCTTGAACCGAATAAATAAACAACAACAAATAA	891

RESULT 6  
Bt412341  
LOCUS Bt412341  
DEFINITION 6029918371 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5147624 5' ,  
mRNA sequence.  
ACCESSION Bt412341  
VERSION Bt412341  
KEYWORDS Bt412341.1 GI:15173264  
SOURCE EST.  
ORGANISM house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 945)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue procurement: Gilbert Smith, Ph.D.









LOCUS	AK013568	754 bp	mRNA	linear	HTC 19-2AN-2002
DEFINITION	Mus musculus adult male hippocampus cDNA, RIKEN full-length insert sequence.				
ACCESSION	AK013568.1 GI:12850979				
VERSION	AK013568				
KEYWORDS	HTC, CAP trapper.				
SOURCE	Mus musculus (strain: C57BL/6J) adult male hippocampus cDNA to mRNA, clone: libb. RIKEN full-length enriched mouse cDNA library clone: 2900022B12.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.				
REFERENCE	Carninci, P. and Hayashizaki, Y.				
AUTHORS	1				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349656				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
AUTHORS	3				
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Wachihi, M., Toned, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076681				
REFERENCE	4				
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Iwawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Imanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadote, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nakado, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staib, F., Nakado, I., Pesole, G., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Botfield, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotmann, L., Hume, D. A., Kamliya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, V., Mombauts, P., Nordone, P., Rling, B., Rongwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Welt, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, R., Kohetsuki, S. and Hayashizaki, Y.				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409 (6821), 685-690 (2001)				
MEDLINE	21085660				
PUBMED	11217851				
REFERENCE	5 (bases 1 to 754)				
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Aizawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Iwawa, M., Kasukawa, T.,				

FEATURES	source
gene	<p>1. 754</p> <p>Location/Qualifiers</p> <p>/organism="Mus musculus"</p> <p>/strain="C57BL/6J"</p> <p>/db_xref="PANTOM.DB:2900022B12"</p> <p>/db_xref="MGD:MG1:1896245"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="2900022B12"</p> <p>/sex="male"</p> <p>/tissue_type="hippocampus"</p> <p>/clone_type="RIKEN full-length enriched mouse cdna library"</p> <p>/der_stage="adult"</p> <p>1. 754</p> <p>/gene="Tmem4"</p> <p>85. 633</p> <p>/gene="Tmem4"</p> <p>/note="data source:MGD, source key:MG1:1928477, evidence:ISS</p> <p>putative</p> <p>transmembrane protein 4"</p> <p>codon_start=1</p> <p>/protein_id="BAB28909.1"</p> <p>/db_xref="GI:12850980"</p> <p>/db_xref="MGD:MG1:1928477"</p> <p>/translation="MKNGMALLALGVLTGTAARRRSDODLHCGCRALVDLELEFIARVDPKRTIOMGSRINPDGOSVVEVPARSAHLTELLEVCDDMKETGEOIDISTRRHKNYRVASRNSESSELDLQIGIRIDSDISGLTFKACESIVEIEYEDLETFESREADNWKDKLSKRTDLCDAHLHSHDEL"</p>
CDS	
BASE COUNT	200 a 175 c 220 g 159 t
ORIGIN	
Query Match	69.8%; Score 745.8; DB 11; Length 754;
Best Local Similarity	99.7%; Pred. No. 3.7e-109;
Matches 747; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
0Y	278 GAGACTCCGGCTGCTTCCAGAGCCTGGGCCACGCGCGCGGCGGGAGCAGAGAGTGA 337
Db	5 GAGAGTCCGGGCTGCTTCCAGAGCCTGGGCCACGCGCGCGGCGGGAGCAGAGTGA 64
0Y	338 GCGACCCGTTACACTAAAGATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGGTCC 397
Db	65 GCGACCCGTTACACTAAAGATGAAGGCTGGGGTGGCTAGCCCTACTTTTGGGGGTCC 124

QY 398 TCGTGGAACTGCTGGGCTCGAAGAGCCAAAGATCTACACTGTGGAGCTTTCAGGGCTC 457  
 |||||||  
 Db 125 TCGTGGAACTGCTGGGCTCGAAGAGCCAAAGATCTACACTGTGGAGCTTTCAGGGCTC 184  
 QY 458 TGGGTGAGATTTAGAGTGGGAAATTTGCCGGGTGGACCCCAAGAGACCAATTCAGATGG 517  
 |||||||  
 Db 185 TGGTGGATGATTTAGAGTGGGAAATTTGCCGGGTGGACCCCAAGAGACCAATTCAGATGG 244  
 QY 518 GATCCTTCGGAATCAATCAGATGGCAGCCAGTCAGTTGTGGAGTACCTTATGCCCGCT 577  
 |||||||  
 Db 245 GATCCTTCGGAATCAATCAGATGGCAGCCAGTCAGTTGTGGAGTACCTTATGCCCGCT 304  
 QY 578 CAGAGGCCACCTCAGAGATTTGCTTGAGAGGTGTGTGACCGAATGAAGAGTACGGGG 637  
 |||||||  
 Db 305 CAGAGGCCACCTCAGAGATTTGCTTGAGAGGTGTGTGACCGAATGAAGAGTACGGGG 364  
 QY 638 AACAGATGACCTTCTCTCCACCCGCAAGAACTACGTACGCTGCTGAGCCGGAATGAG 697  
 |||||||  
 Db 365 AACAGATGACCTTCTCTCCACCCGCAAGAACTACGTACGCTGCTGAGCCGGAATGAG 424  
 QY 698 AATCCAGTAACTAGACTTACAGGGCATCCGAATTCAGATATACGCGGCAACCTCA 757  
 |||||||  
 Db 425 AATCCAGTAACTAGACTTACAGGGCATCCGAATTCAGATATACGCGGCAACCTCA 484  
 QY 758 AGTTTGGCTGTGAGAGCTTTGTGGAAGAATACAGAGATGAGCTTATCCAAATTTCTTCCA 817  
 |||||||  
 Db 485 AGTTTGGCTGTGAGAGCTTTGTGGAAGAATACAGAGATGAGCTTATCCAAATTTCTTCCA 544  
 QY 818 GAGAGGCTGACAGCTTTAAAGACAAACTTTGACAGTAAAGGACAGATATATGACATG 877  
 |||||||  
 Db 545 GAGAGGCTGACAGCTTTAAAGACAAACTTTGACAGTAAAGGACAGATATATGACATG 604  
 QY 878 CCTGTACACAGATCTACAGATAGTATGATCACTGAGACAGACCTTACACCAAGCT 937  
 |||||||  
 Db 605 CCTGTACACAGATCTACAGATAGTATGATCACTGAGACAGACCTTACACCAAGCT 664  
 QY 938 GATGAACACCCCAAGAGGGGAGATGGCAGATGCTTTTATATACGTTTTTATAG 997  
 |||||||  
 Db 665 GATGAACACCCCAAGAGGGGAGATGGCAGATGCTTTTATATACGTTTTTATAG 724  
 QY 998 AATGAACTGAAAAAACTCTTGAACCG 1026  
 |||||||  
 Db 725 AATGAACTGAAAAAACTCTTGAACCG 753  
 RESULT 9  
 BF608664  
 LOCUS  
 DEFINITION 802 bp mRNA linear EST 01-APR-2001  
 M1.001606 Mouse 9-day fetus cdna library ICRP522 Mus musculus  
 ACCESSION  
 BF608664  
 VERSION  
 BF608664.1 GI:13505156  
 KEYWORDS  
 EST.  
 SOURCE  
 house mouse.  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;  
 REFERENCE  
 1 (bases 1 to 802)  
 Yahyawi, M., Hemmig, S., Neidhardt, L., Radelof, U., Hermann, B. G.,  
 Lehrich, H. and O'Brien, J.  
 TITLE  
 Detection of a high number of novel genes in a 9-day mouse embryo  
 JOURNAL  
 cdna library normalised by oligonucleotide fingerprinting  
 COMMENT  
 Unpublished (2001)  
 Contact: Hennig S  
 Laboratory 123, dept. Lehrach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr. 63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1612  
 Fax: +49 30 8413 1380  
 Email: hennig@molgen.mpg.de  
 EST's are made from clones being representatives of clone clusters.  
 Clone clusters were calculated from oligonucleotide fingerprints.

FEATURES  
 source  
 location/qualifiers  
 1..802  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="ICRP522B2440"  
 /clone\_lib="Mouse 9-day fetus cdna library ICRP522"  
 /issue\_type="whole embryo"  
 /dev\_stage="embryonic 9-day"  
 /lab\_host="E.coli, XLI blue"  
 /note="Vector: pSVSPORT1; Site\_1: NotI; Site\_2: SalI;  
 Library preparation by oligo-dt priming of RNA. Clones can  
 be ordered from the Resource Center in Berlin,  
 http://www.rzpd.de."

BASE COUNT 193 a 206 c 243 g 159 t 1 others  
 ORIGIN  
 Query Match 69.7%; Score 744.6; DB 12; Length 802;  
 Best Local Similarity 99.3%; Pred. No. 5.5e-109;  
 Matches 747; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 169 GGTCACTGGCAACCTGAACTGGCGGAGCCGAGCCCTTGAGAAACCGGACACCC 228  
 |||||||  
 Db 51 GGTCACTGGCAACCTGAACTGGCGGAGCCGAGCCCTTGAGAAACCGGACACCC 110  
 QY 229 CCGTTTTCACAGCAGCTGTGAGGCTTGGGACCCAGAAACAAAGTGAAGTCCGGC 288  
 |||||||  
 Db 111 CCGTTTTCACAGCAGCTGTGAGGCTTGGGACCCAGAAACAAAGTGAAGTCCGGC 170  
 QY 289 TCGTTTCCAGAGCCTTGCGCAACGCGCGCGCGCTGTGGAGCAGAGGTGGAGCACCCTGTT 348  
 |||||||  
 Db 171 TCGTTTCCAGAGCCTTGCGCAACGCGCGCGCGCTGTGGAGCAGAGGTGGAGCACCCTGTT 230  
 QY 349 ACACATAAGATGAAGAGCTGGGGTGGCTAGCCCTACTTTTGGGGGCTCTCTGGGAAT 408  
 |||||||  
 Db 231 ACACATAAGATGAAGAGCTGGGGTGGCTAGCCCTACTTTTGGGGGCTCTCTGGGAAT 290  
 QY 409 GCTGGGCTCGAAGAGGACCAAGATCTACAGTGTGAGGCTTCAAGGCGCTGTGGATGA 468  
 |||||||  
 Db 291 GCTGGGCTCGAAGAGGACCAAGATCTACAGTGTGAGGCTTCAAGGCGCTGTGGATGA 350  
 QY 469 TTAGAGTGGAAATTTGCCCGGTGGACCCCAAGAACCAATTCAGATGGATCTCTCCGA 528  
 |||||||  
 Db 351 TTAGAGTGGAAATTTGCCCGGTGGACCCCAAGAACCAATTCAGATGGATCTCTCCGA 410  
 QY 529 ATCAATCCAGATGGCAGCAGTCACTTGTGAGAGTACTTATGCCGCTCAGAGGCCAC 588  
 |||||||  
 Db 411 ATCAATCCAGATGGCAGCAGTCACTTGTGAGAGTACTTATGCCGCTCAGAGGCCAC 470  
 QY 589 CTCACAGATGCTTGAAGAGGTGTGACCAAGATGAAGAGTACGGGAGACAGATTGAC 648  
 |||||||  
 Db 471 CTCACAGATGCTTGAAGAGGTGTGACCAAGATGAAGAGTACGGGAGACAGATTGAC 530  
 QY 649 CTTTCTACCAACCCGCAAGATACGTACGCTGTGAGCCGGAATGGAATTCAGATGA 708  
 |||||||  
 Db 531 CTTTCTACCAACCCGCAAGATACGTACGCTGTGAGCCGGAATGGAATTCAGATGA 590  
 QY 709 CTAGACTTACAGGGCATCCGAATTTGACTCAGATATACAGGCGACCCCTCAAGTTGGCTGT 768  
 |||||||  
 Db 591 CTAGACTTACAGGGCATCCGAATTTGACTCAGATATACAGGCGACCCCTCAAGTTGGCTGT 650  
 QY 769 GAGAGCATTTGGAAGATATAGAGATGAGTATTCGAATCTCTCTCCAGAGAGCTGAC 828  
 |||||||  
 Db 651 GAGAGCATTTGGAAGATATAGAGATGAGTATTCGAATCTCTCTCCAGAGAGCTGAC 710  
 QY 829 AACGTTAAAGACAACTTTCAGTAAAGGAGCAGATCTATGTGACATGCCCTGCACAGA 888  
 |||||||  
 Db 711 CACGTTAAAGACAACTTTCAGTAAAGGAGCAGATCTATGTGACATGCCCTGCACAGA 770







DB 771 CTTTATATTTAGTTTATGAAATGAACTGAAAAAACTCTGAAAAACCGAAAAA 830  
 OY 1035 AAAAAAAAAA 1046  
 DB 831 AAGGAGAAAA 842

RESULT 14  
 BI411532 709 bp mRNA linear EST 14-AUG-2001  
 LOCUS 602965126F1.NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5120574 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI411532.1 GI:15172455  
 VERSION BI411532.1  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 709)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L14M11294 row: c column: 07  
 High quality sequence start: 4  
 High quality sequence stop: 701.

# FEATURES source

1. 709  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5120574"  
 /clone\_1lb="NCI-CGAP\_Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pRTT3-Pac (Pharmacia) with a  
 modified polylinker; Site:1: NotI; Site:2: EcoRI; 1st  
 strand cDNA was prepared from mRNA digested with Not I and  
 Eco RI sites of the modified pRTT3 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 191 a 170 c 214 g 132 t 2 others  
 ORIGIN

Query Match 64.2% Score 686; DB 13; Length 709;  
 Best Local Similarity 98.9%; Pred. No. 1.2e-99;  
 Matches 700; Conservative 0; Mismatches 7; Indels 1; Gaps 1;  
 OY 256 TGGAGCCACAGAAAAAGTAGAGTCGCGCTCTTCCAGAGCTGGCCACGGCGG 315  
 DB 3 TGGAGCCACAGAAAAAGTAGAGTCGCGCTCTTCCAGAGCTGGCCACGGCGG 62  
 OY 316 CGGCGGTGGAGACAGAGTGGAGCAGCCCTGTACACTTAAGATGAAGAGCTGGGTTGG 375  
 DB 63 CGGCGGTGGAGACAGAGTGGAGCAGCCCTGTACACTTAAGATGAAGAGCTGGGTTGG 121  
 OY 376 CTAGCCCTACTTTTGGGGTCTGCTGGAACTGCTGGGCTGAAGAGCAAGATCTA 435  
 DB 122 CTAGCCCTACTTTTGGGGTCTGCTGGAACTGCTGGGCTGAAGAGCAAGATCTA 181  
 OY 436 CACTGTGAGCTTTCAGAGGCTCTGTGTGATGATTAAGTGGGAAATTCGCCGCTGGAC 495  
 DB 182 CACTGTGAGCTTTCAGAGGCTCTGTGTGATGATTAAGTGGGAAATTCGCCGCTGGAC 241  
 OY 496 CCCAGAGACCAATTCAGATGGGATCTTCGCAATCAATCCAGATGGCAGCCAGTCAATT 555  
 DB 242 CCCAGAGACCAATTCAGATGGGATCTTCGCAATCAATCCAGATGGCAGCCAGTCAATT 301  
 OY 556 GTGGAGGTACTTTATGCCCCCTCAGAGGCCACCTCAGAGATGGTGTAGAGAGGTGTGT 615  
 DB 302 GTGGAGGTACTTTATGCCCCCTCAGAGGCCACCTCAGAGATGGTGTAGAGAGGTGTGT 361  
 OY 616 GACCGAATGAGAGTAGAGGAGGAGACAGATGACCTTTCACCCACCGCAAGAACTACGTA 675  
 DB 362 GACCGAATGAGAGTAGAGGAGGAGACAGATGACCTTTCACCCACCGCAAGAACTACGTA 421  
 OY 676 CGCGTGTAGCCGGAATGAGAAATTCAGATGAATGACTTACAGGGCATCCGAATTGAC 735  
 DB 422 CGCGTGTAGCCGGAATGAGAAATTCAGATGAATGACTTACAGGGCATCCGAATTGAC 481  
 OY 736 TCAGATATCAGCGGCACCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGAT 795  
 DB 482 TCAGATATCAGCGGCACCTCAAGTTTGGCTGTGAGAGCATTTGGAGAAATACGAGAT 541  
 OY 796 GAGCTTATGAAATTTCTTCAGAGAGGCTGACAACTTAAGACAACTTTGCACTAAG 855  
 DB 542 GAGCTTATGAAATTTCTTCAGAGAGGCTGACAACTTAAGACAACTTTGCACTAAG 601  
 OY 856 CGGACAGATCTATGTGACCATATGCTGACACAGATCTCAGATGAGTATGATCTGCA 915  
 DB 602 CGGACAGATCTATGTGACCATATGCTGACACAGATCTCAGATGAGTATGATCTGCA 661  
 OY 916 GCAAGCAGCTACACCAAGCTGATGAGAACACCCCGAGAGGGGAGGA 963  
 DB 662 GCAAGCAGCTACACCAAGCTGATGAGAACACCCCGAGAGGGGAGGA 709

RESULT 15  
 BQ963382 894 bp mRNA linear EST 21-AUG-2002  
 LOCUS AGENCOURT.10050169 NIH\_MGC.134 Mus musculus cDNA clone  
 DEFINITION IMAGE:6508512 5', mRNA sequence.  
 ACCESSION BQ963382  
 VERSION BQ963382.1 GI:22378860  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 894)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Dr. David Rowe  
 cDNA Library Preparation: Invitrogen Corp  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L14M14074 row: f column: 01  
 High quality sequence stop: 528.

## FEATURES source

1. 894  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6508512"  
 /clone\_1lb="NIH-MGC.134"  
 /tissue\_type="undifferentiated limb"



```
/lab_host="DH10B (phage-resistant)"
/Note="Vector: PCMV-SPORT6.1.cdb; Site.1: EcoRV; Site.2:
NotI; Cloned unidirectionally. Primer: Oligo dT. Average
Insert size 1.7 kb. Constructed by Reagen, Invitrogen
Corp. Note: this is a NIH MGC Library."
BASE COUNT      225 a      232 c      268 g      168 t
ORIGIN           1 others
```

```
Query Match      61.2%; Score 654.4; DB 14; Length 894;
Best Local Similarity 99.6%; Pred. No. 1e-94;
Matches 666; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 129 CGCTTAGTGCCCTAGTTAGTGGGGAACTGGCCAGCGGTCACTGAGCCCTGAAC 188
    |||||||
DB 6 CGCTTAGTGCCCTAGTTAGTGGGGAACTGGCCAGCGGTCACTGAGCCCTGAAC 65
OY 189 CTGGCGGGAGCCGAGCCGCTGTGAGAAAGCCGAGACGCCCGTTTTCACGACGCTG 248
    |||||||
DB 66 CTGGCGGGAGCCGAGCCGCTGTGAGAAAGCCGAGACGCCCGTTTTCACGACGCTG 125
OY 249 CTAGGGTTGGGACCAAGAAAGTGAAGTCCGGCTGTTTCCAGAGCCCTGGGCC 308
    |||||||
DB 126 CTAGGGTTGGGACCAAGAAAGTGAAGTCCGGCTGTTTCCAGAGCCCTGGGCC 185
OY 309 ACGGCGGCGCCGCTGGGAGCAGAGGTGAGCGACCCCTGTTACACTAAAGATGAAGGCTG 368
    |||||||
DB 186 ACGGCGGCGCCGCTGGGAGCAGAGGTGAGCGACCCCTGTTACACTAAAGATGAAGGCTG 245
OY 369 GGGTTGGCTAGCCCTACTTTTGGGGGCTCTCTGTGGAACTGCTGGGCTCGAAGAGCCA 428
    |||||||
DB 246 GGGTTGGCTAGCCCTACTTTTGGGGGCTCTCTGTGGAACTGCTGGGCTCGAAGAGCCA 305
OY 429 AGATCTACACTGTGAGGCTTCAGAGGCTCTGTGGATGAATTAGAGTGGGAATTGCCCG 488
    |||||||
DB 306 AGATCTACACTGTGAGGCTTCAGAGGCTCTGTGGATGAATTAGAGTGGGAATTGCCCG 365
OY 489 CGTGAGACCCCAAGAAAGACATTCAAGATGGATCCTTCGATCAATCCAGATGCGACCA 548
    |||||||
DB 366 CGTGAGACCCCAAGAAAGACATTCAAGATGGATCCTTCGATCAATCCAGATGCGACCA 425
OY 549 GTCAATTGTGAGGTACCTTATGCCCGCTCAGAGGCCACCTCAGAGATTGCTTGAGA 608
    |||||||
DB 426 GTCAATTGTGAGGTACCTTATGCCCGCTCAGAGGCCACCTCAGAGATTGCTTGAGA 485
OY 609 GGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCCCTTACCCACCGCAAGAA 668
    |||||||
DB 486 GGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCCCTTACCCACCGCAAGAA 545
OY 669 CTAGCTACGGGTGCTGAGCCCGGAATGGAATCCAGTGAATCTAGACTTACAGGCGCATCCG 728
    |||||||
DB 546 CTAGCTACGGGTGCTGAGCCCGGAATGGAATCCAGTGAATCTAGACTTACAGGCGCATCCG 605
OY 729 AATTGACTCAGATATCAGCGCACCCCTCAAGTTGGGTGAG -AGCATTTGGAAAGAA 787
    |||||||
DB 606 AATTGACTCAGATATCAGCGCACCCCTCAAGTTGGGTGAG -AGCATTTGGAAAGAA 665
OY 788 ACGAGGATG 796
    |||||||
DB 666 ACGAGGATG 674
```

Search completed: March 30, 2003, 20:56:53  
Job time : 1531.26 secs





GenCore version 5.1.4.D5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:43:09 ; Search time 1166.6 Seconds

(without alignments)  
2526.643 Million cell updates/sec

Title: US-10-082-502-19  
Perfect score: 959  
Sequence: 1 MKGGMALLIGVLTGTAWA.....LCSTRDLCDAHRSDEL 182

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL-frame+P2n.model -DEV-x1p  
-Q/cgn2\_1/USPTO.spool/US10082502/runat.24032003\_135100\_6599/app-query.fasta.1.1308  
-DB-EST -QPMT-fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-Blosum62 -TRANS-human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-USER=us10082502 -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-NO\_XLPRY -NO\_MMAP -LARGEOUTERY -NES-SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_hic3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_juv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	100.0	754	11 AK013568	AK013568 Mus muscu
2	959	100.0	772	11 AK007914	AK007914 Mus muscu
3	959	100.0	808	13 B1408525	B1408525 602964995
4	959	100.0	843	13 B1412539	B1412539 602990995
5	959	100.0	853	11 AK013014	AK013014 Mus muscu
6	959	100.0	1255	11 AK019927	AK019927 Mus muscu
7	947	98.7	802	12 BF608664	BF608664 MY1.00160
8	946	98.6	891	13 B1408381	B1408381 602964013
9	943	98.3	724	14 BQ210416	BQ210416 UT-R-DY1-
10	938	97.8	588	14 BM832828	BM832828 K-EST0107
11	938	97.8	591	14 BM832888	BM832888 K-EST0107
12	938	97.8	611	13 B1712899	B1712899 1d98f09.Y
13	938	97.8	697	12 BF344334	BF344334 602014708
14	938	97.8	719	12 BE733900	BE733900 601568370
15	938	97.8	742	12 BG824549	BG824549 602728454
16	938	97.8	745	12 BE791763	BE791763 601581586
17	938	97.8	1117	13 BM552907	BM552907 AGENCOURT
18	932	97.2	880	14 BQ421827	BQ421827 AGENCOURT
19	930	97.0	783	13 B1550436	B1550436 603132905
20	925	96.5	868	14 BQ421059	BQ421059 AGENCOURT
21	923	96.2	791	12 BG704443	BG704443 602688633
22	920	95.9	856	13 B1454180	B1454180 603170602
23	920	95.9	992	13 B1456021	B1456021 603170602
24	917	95.6	773	13 B1550162	B1550162 603193045
25	913	95.2	709	13 B1411532	B1411532 602965126
26	912	95.1	818	13 B1763950	B1763950 603049829
27	912	95.1	878	12 BE870315	BE870315 601447464
28	911	95.0	887	13 B1762759	B1762759 603048045
29	910	94.9	665	10 AW919569	AW919569 EST350873
30	906	94.5	720	13 B1909759	B1909759 603070841
31	905	94.4	673	14 BQ668530	BQ668530 AGENCOURT
32	898	93.6	783	13 B1763784	B1763784 603049729
33	898	93.6	914	12 BE799445	BE799445 601889289
34	891	92.9	765	12 BF312281	BF312281 601889289
35	886	92.4	678	10 BE309953	BE309953 601091903
36	884	92.2	856	12 BF965088	BF965088 602268856
37	878	91.6	751	12 BE742698	BE742698 601574753
38	874	91.1	657	14 BQ211867	BQ211867 UT-R-DY1-
39	874	91.1	945	13 B1412341	B1412341 602991837
40	870	90.7	867	13 BG973691	BG973691 602844540
41	870	90.7	885	14 BQ920767	BQ920767 AGENCOURT
42	870	90.7	963	14 BQ929020	BQ929020 AGENCOURT
43	869	90.6	674	14 BQ202181	BQ202181 UT-R-EB1-
44	868	90.5	797	13 B1762927	B1762927 603047987
45	866	90.3	549	10 BE198865	BE198865 u983d11.Y

## ALIGNMENTS

RESULT 1  
AK013568  
LOCUS  
DEFINITION  
AK013568 754 bp mRNA linear HIC 19-JAN-2002  
Mus musculus adult male hippocampus cDNA, RIKEN full-length  
enriched library, clone:2900022B12:transmembrane protein 4, full  
insert sequence.  
ACCESSION  
AK013568  
VERSION  
AK013568.1 GI:12850979  
KEYWORDS  
HIC; CAP trapper;  
SOURCE  
Mus musculus (strain:C57BL/6J) adult male hippocampus cDNA to mRNA,  
clone:2900022B12.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	AUTHORS	TITLE	JOURNAL	PUBLISHED
1	Carninci, P. and Hayashizaki, Y.	High-efficiency full-length cDNA cloning	Mech. Enzymol.	303, 19-44 (1993)
2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Genome Res.	10 (10), 1617-1630 (2000)
3	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Genome Res. 10 (11), 1757-1771 (2000)		
4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Alizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Eletschmann, W., Gaasterland, T., Glass, C., King, B., Kochiya, H., Kuehl, P., Lewis, S., Matsuo, Y., Staudt, I., Pesole, G., Quackenbush, J., Schiraldi, L., Staudt, I., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, C., Hume, D. A., Kamaya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilmberg, J., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA collection	Nature	409 (6821), 685-690 (2001)
5	(bases 1 to 754)			
6	Adachi, J., Alizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bul, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imoto, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M., Koyama, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohtsuka, K., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Schiraldi, L., Shibata, K., Shibata, Y., Shihagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	Direct Submission		
7	Submitted (10-JUL-2000)	Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.resc.riken.go.jp/, Tel: 81-45-503-9222,		



transmembrane protein 4"  
/codon\_start=1  
/protein\_id="BAB25346.1"  
/db\_xref="GI:12841774"  
/db\_xref="MGI:1928477"  
/translation="MKMGWLLLLGLVLTAMARRSODLHGCARATVDELEMETAR  
VDPKTIQMGSRFNPDSQSYVEVPYPASEAHLLELVEYCDRKEVEQDIPSTHR  
KATRVVSRNGSSSELDQGIINDSIDISGLFACBSLYEETLEDELEFFSRKADNVK  
DKLSKRDLCDHALHRSHDEL"

BASE COUNT 210 a 180 c 225 g 157 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.55e-108 Length: 772  
Score: 959.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-082-502-19 (1-182) x AK007914 (1-772)

QY 1 MetlySGlyTRpGlyTRpleuAlaleuLeuGlyValleuLeuGlyThrAlaTRpAla 20  
|||||  
Db 104 ATGAAGGCTGGGGTGGCTAGCCCTACTTTGGGGGCTCCTGGGAAGCTGCTGGGCT 163  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTRp 40  
|||||  
Db 164 CGAAGAGCCAGATGCTACCTGTGAGGCTTGCAAGGCTCTGTGTGATGATTAAGATGG 223  
QY 41 GluIleAlaArgValAspProlylsThrIleGlnMetClySerPheArgIleAsnPro 60  
|||||  
Db 224 GAATATGCCCGCTGGAGCCCAAGAGACCATTCAGATGGATCCCTCCGAATCAATCCA 283  
QY 61 AspGlySerGlnSerValAlaGluValProtyrAlaArgSerGlnAlaHisLeuThrGlu 80  
|||||  
Db 284 GATGGACGACGACGATGTGGAGGTACCTTATGCCCGCTCAAGGCCCACTCACAGAG 343  
QY 81 LeuLeuGluGluValAlaCysAspArgMetlyGlyTrpGlyGluGlnIleAspProSerThr 100  
|||||  
Db 344 TTGCTTGAGAGAGGTGTGTGACCGAATGAAGAGTACGGGACAGATTCACCTTCTAC 403  
QY 101 HisArgLyAsnTrpValAlaArgValAlaSerArgAsnGlyIleSerSerGluLeuAspLeu 120  
|||||  
Db 404 CACCGCAAGACTACGTACGCTGCTGAGCCGGAATGAGAAATCCAGTGAAGACTTA 463  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLyPheAlaCysGluSerIle 140  
|||||  
Db 464 CAGGCGATCGAATTCAGATTCAGATTCAGGCGCACCTCAAGTTTGGTGTGAGACATT 523  
QY 141 ValGluGluTrpGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
|||||  
Db 524 GTGGGAAGATACAGAGATGAGCTTATCGAATTTCTTCCAGAGAGGCTGCACAACGTTAA 583  
QY 161 AspLyLeuLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
|||||  
Db 584 GCAAAACCTTGCAATGAGGAGCAGATCATGTGACCATCCCTGCACAGATCTCAGAT 643  
QY 181 GluLeu 182  
|||||  
Db 644 GAGCTA 649

RESULT 3  
BI408525 808 bp mRNA linear EST 14-AUG-2001  
LOCUS 60296495F1 NCI\_GAP\_Lu33 Mus musculus cDNA clone IMAGE:5120471 5',  
DEFINITION mRNA sequence.  
ACCESSION BI408525  
VERSION BI408525.1 GI:15169448  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 808)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM1293 row: n column: 24  
High quality sequence start: 4  
High quality sequence stop: 768.  
Location/Qualifiers  
1. 808  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5120471"  
/clone\_1lb="NCI-CGAP\_Lu33"  
/tissue\_type="pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCGGCTGTGTTTATTTTATTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 224 a 181 c 237 g 166 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.93e-108 Length: 808  
Score: 959.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-082-502-19 (1-182) x BI408525 (1-808)

QY 1 MetlySGlyTRpGlyTRpleuAlaleuLeuGlyValleuLeuGlyThrAlaTRpAla 20  
|||||  
Db 119 ATGAAGGCTGGGGTGGCTAGCCCTACTTTGGGGGCTCCTGGGAAGCTGCTGGGCT 178  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTRp 40  
|||||  
Db 179 CGAAGAGCCAGATGCTACCTGTGAGGCTTGCAAGGCTCTGTGTGATGATTAAGATGG 238  
QY 41 GluIleAlaArgValAspProlylsThrIleGlnMetClySerPheArgIleAsnPro 60  
|||||  
Db 239 GAATATGCCCGCTGGAGCCCAAGAGACCATTCAGATGGATCCCTCCGAATCAATCCA 298  
QY 61 AspGlySerGlnSerValAlaGluValProtyrAlaArgSerGlnAlaHisLeuThrGlu 80  
|||||  
Db 299 GATGGACGACGACGATGTGGAGGTACCTTATGCCCGCTCAAGGCCCACTCACAGAG 358  
QY 81 LeuLeuGluGluValAlaCysAspArgMetlyGlyTrpGlyGluGlnIleAspProSerThr 100  
|||||  
Db 359 TTGCTTGAGAGAGGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCTTCTAC 418  
QY 101 HisArgLyAsnTrpValAlaArgValAlaSerArgAsnGlyIleSerSerGluLeuAspLeu 120  
|||||  
Db 419 CACCGCAAGACTACGTACGCTGTGAGCGGGAATGGAATCCAGTGAAGTAACTTA 478

QY 121 GINGLYTlearglleasperaspiileserGlythrLeuLysPhealaCysgluSerlle 140  
DB 479 CAGGCGATCCGAAATGACTACAGATATCAGCGACCCCTCAAGTTTCGCTGAGACATT 538  
QY 141 ValGluGluTyrGluaspGluLeuLeuileGluPhePheSerArgGluAlaaspasVallys 160  
DB 539 GTGGAAAGATACAGAGATAGCTTATCGAATCTTCTCCAGACAGGCTGACAACTTTAA 598  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCyaspPheAlaLeuLysArgSerHisasp 180  
DB 599 GACAAACTTTCGAGTAAGCGGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGAT 658  
QY 181 GluLeu 182  
DB 659 GAACCTT 664  
RESULT 4  
LOCUS B1412539 843 bp mRNA linear EST 14-AUG-2001  
DEFINITION 60290995F1 NC1\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5146967 5',  
mRNA sequence.  
ACCESSION B1412539  
VERSION B1412539.1 GI:15173462  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/  
AUTHORS 1 (bases 1 to 843)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaab@r-email.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldi, Ph.D.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NC1-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
place: LLM11362 row: n column: 24  
High quality sequence start: 38  
High quality sequence stop: 824.  
Location/Qualifiers  
FEATURES  
source 1..843  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5146967"  
/clone\_1lb="NC1\_CGAP\_Lu33"  
/tissue\_type="pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pTZ19-3-Pac (Pharmacia) with a  
modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand cDNA was prepared from mRNA obtained from pooled  
lung tumors with a Not I - oligo(dT) primer [5'  
TGTTCACATCTGAGTGGAGCGGCCGCTCTGTTTCTTTTCTTTTCTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified pTZ19 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldi. "

Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0  
US-10-082-502-19 (1-182) x B1412539 (1-843)  
QY 1 MetLysGlyTyrPGLYThrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
DB 154 ATCAAAAGCTGGGGTGGCTGAGCTTACAGCTTGTGAGGCTGTGGATTAATTAAGATGG 213  
QY 21 ArgArgSerGluAspPheHisCysGlyValaCysArgAlaLeuValAspLeuLeuGluTrp 40  
DB 214 CGAAGAGCGCAAGATCTACAGCTGTGAGCTTGCAGGCGCTGTGGATTAATTAAGATGG 273  
QY 41 GluLeuAlaArgValaAspProLysLysThrileGlnMetGlySerPheArgLysAspPro 60  
DB 274 GAATGTGCCCGGGTGGAGCCCAAGAGACCATTCAGATGGATCCCTTCGAAATCAATCA 333  
QY 61 AspGlySerGlnSerValaValaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 334 GATGGCAGCCAGTCACTTGTGAGTACCTTATGCCCCCTCAGAGGCCACCTTCACAGAG 393  
QY 81 LeuLeuGluGluValaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
DB 394 TTGCTTGAGGAGGTGTGTGACCAATGAAGAGTACGGGGAACAGATTACCTTCTTACC 453  
QY 101 HisArgLysAsnTyrValaArgValaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 454 CACCGCAAGACACTAGCTAGCGCTGAGCCGGAATGAGAAATCCAGTAACTAGACTTA 513  
QY 121 GINGLYTlearglleasperaspiileserGlythrLeuLysPhealaCysgluSerlle 140  
DB 514 CAGGCGATCCGAAATGACTACAGATATCAGCGACCCCTCAAGTTTCGCTGAGACATT 573  
QY 141 ValGluGluTyrGluaspGluLeuLeuileGluPhePheSerArgGluAlaaspasVallys 160  
DB 574 GTGGAAAGATACAGAGATAGCTTATCGAATCTTCTCCAGAGAGGCTCACAACGTTAA 633  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCyaspPheAlaLeuLysArgSerHisasp 180  
DB 634 GACAAACTTTCGAGTAAGCGGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGAT 693  
QY 181 GluLeu 182  
DB 694 GAACCTT 699  
RESULT 5  
LOCUS AK013014 853 bp mRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
enriched library, clone:2810406L24:transmembrane protein 4, full  
insert sequence.  
ACCESSION AK013014  
VERSION AK013014.1 GI:12850114  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA,  
clone:2810406L24.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE Carninci, P. and Hayashizaki, Y.  
AUTHORS High-efficiency full-length cDNA cloning  
TITLE Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL 99279253  
MEDLINE 10349636  
PUBMED 10349636  
REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
AUTHORS Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes  
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)  
20499374





DB	634	GENAGAAATACGAGGATGAGCTTATGCAATTCCTCCAGAGAGCGACAGCAACTTAA	693
OY	161	Asp1y5Ieucy5set1yarg7rthAsp1eucy5Asp1h1Ala1euh1Iarg5er1asp	180
DB	694	GACAAATCTTTCGATGACGCGGACAGATCTATGTGACATGCGCTCCACAGATTCACGAT	753
OY	181	Glueu	182
DB	754	GNCTA	759
RESULT 6	AK019927	1255 bp	HTC 19-JAN-2002
LOCUS	AK019927		
DEFINITION	Mus musculus adult male pituitary gland cDNA, RIKEN full-length		
	enriched library, clone:5330432A10:transmembrane protein 4, full		
	insert sequence.		
ACCESSION	AK019927		
VERSION	AK019927.1	GI:12860332	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male pituitary gland cDNA to		
	cDNA, clone:lib:RIKEN full-length enriched mouse cDNA library		
	clone:5330432A10.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	1	Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning	
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)	
PUBMED		99279253	
REFERENCE		10349636	
AUTHORS	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
TITLE		Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
JOURNAL		Normalization and subtraction of cap-trapper-selected cDNAs to	
PUBMED		prepare full-length cDNA libraries for rapid discovery of new genes	
REFERENCE		Genome Res. 10 (10), 1617-1630 (2000)	
AUTHORS	3		
TITLE		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
JOURNAL		Kono, H., Akiyama, J., Nishik, K., Kitsumi, T., Tashiro, H., Itoh, M.,	
PUBMED		Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,	
REFERENCE		Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K.,	
AUTHORS		Fujiwaka, K., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watabiki, M.,	
TITLE		Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuno, S., Kawai, J.,	
JOURNAL		Okazaki, Y., Muramatsu, M., Inoue, Y., Kiru, A. and Hayashizaki, Y.	
PUBMED		RIKEN integrated sequence analysis (RISA) system--384-format	
REFERENCE		sequencing pipeline with 384 multichannel sequencer	
AUTHORS	4	Genome Res. 10 (11), 1757-1771 (2000)	
TITLE		11076861	
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			

JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PubMed	11217851
AUTHORS	5 (bases 1 to 1255) Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Cantucci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kuzihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Schirml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physiological and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGATTCGACGATTAATTAATATATCCCCCCCCC 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 3.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATTCGACGATTAATTAATATATCCCCCCCCC 3']. cDNA was amplified by PCR with BamHI and XhoI. Vector: a modified pBlueScript KS(+) after bulk excision from lambda PhiC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.
FEATURES	Location/Qualifiers 1..1255 /oranism="Mus musculus" /strain="C57BL/6J" /db_xref="FANTOM_DB:5330432A10" /db_xref="MGD:MGI:1893093" /db_xref="taxon:10090" /clone="5330432A10" /sex="male" /tissue_type="pituitary gland" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 1..1255 /gene="Tmem4" 215..763 /gene="Tmem4" /note="data source:MGD, source key:MGI:1928477, evidence:ISS putative transmembrane protein 4" /codon_start=1 /protein_id="BAB31921.1" /db_xref="GI:12860333" /db_xref="MGD:MGI:1928477" /translation="MGKGMVLLGLVLTGTAMARSDDLHGACRALVDLEWEIARVDPKTIOMGSEFRINPDGSSVVEVPARSANHTELEEVCDMKKEVGEOIDPSSTRKNRYRARSNESESLDIOGIRIDSDISGTLKFACESEIVEYEDLEIFFSREADNRYKDKLSKRTDLCDAHLHSHDEL"
gene	
CDS	
BASE COUNT	309 a 293 c 354 g 299 t
ORIGIN	

Alignment Scores:

Pred. No.:	1.12e-107	Length:	1255
Score:	959.00	Matches:	182
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	11	Gaps:	0

US-10-082-502-19 (1-182) x AK019927 (1-1255)

QY 1 MetLysGlyTyrPglYTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
|||||  
DB 215 ATGAAAGGCTGGGGTGGCTGAGCCCTACTTTGGGGGCTGCTGGGAAGCTGCTGGCT 274

QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
|||||  
DB 275 CGAAGAGCCAAAGATCTACAGCTGAGAGCTTGAGAGGCTCTGTGATGATTAAGTGG 334

QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60  
|||||  
DB 335 GAATATGCCCGGGTGGAGCCCAAGAACCATTCAGATGGATCCCTCCGAATCAATCCA 394

QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
|||||  
DB 395 GATGGCAGCAGCTCACTTGGAGGTACTTATGCCGCTCAGAGGCCACCTCAGAGAG 454

QY 81 LeuLeuGluGluValAlaCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
|||||  
DB 455 TTGCTTGAAGAGGTGTGTGACCGAATGAAAGAGTACGGGGAACAGATTACCCCTTACC 514

QY 101 HisArgGlyAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
|||||  
DB 515 CACCGCAAGACTACGATACGCTGCGTGGAGCCGGAATGGGAATCCAGTACTAGACTTA 574

QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
|||||  
DB 575 CAGGCGATCCGAATTCACATATACACGAGGACCCCTTAAGTTTCGTGTGAGACATT 634

QY 141 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 160  
|||||  
DB 635 GTGGAAGATACGAGAGTACGCTTATCGAATCTCTCCAGAGAGCTCACAACGTTAA 694

QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
|||||  
DB 695 GACAAACTTGGAGTAAAGGAGACAGATCTATGTGACCAATGCCCTGACAGATCTCAGAT 754

QY 181 GluLeu 182  
|||||  
DB 755 GAGCTA 760

RESULT 7  
BF608664 802 bp mRNA linear EST 01-APR-2001  
LOCUS MYL1001606 Mouse 9-day fetus cDNA library ICRF522 Mus musculus  
DEFINITION cDNA clone ICRF522B2440 5', mRNA sequence.  
ACCESSION BF608664  
VERSION BF608664.1 GI:13505156  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 802)  
AUTHORS Yabuyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B. G.,  
Lehrach, H. and O'Brien, J.  
TITLE Detection of a high number of novel genes in a 9-day mouse embryo  
cDNA library normalised by oligonucleotide fingerprinting  
JOURNAL Unpublished (2001)  
COMMENT Contact: Hennig S  
Laboratory 123, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inhestr. 63-73, D-14195 Berlin, Germany

Alignment Scores:

Pred. No.:	1.78e-106	Length:	802
Score:	947.00	Matches:	180
Percent Similarity:	99.45%	Conservative:	1
Best Local Similarity:	98.90%	Mismatches:	0
Query Match:	98.75%	Indels:	0
DB:	12	Gaps:	0

US-10-082-502-19 (1-182) x BF608664 (1-802)

QY 1 MetLysGlyTyrPglYTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
|||||  
DB 240 ATGAAAGGCTGGGGTGGCTGAGCCCTACTTTGGGGGCTGCTGGGAAGCTGCTGGCT 299

QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
|||||  
DB 300 CGAAGAGCCAAAGATCTACAGCTGAGAGCTTGAGAGGCTCTGTGATGATTAAGTGG 359

QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 60  
|||||  
DB 360 GAATATGCCCGGGTGGAGCCCAAGAACCATTCAGATGGATCCCTCCGAATCAATCCA 419

QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
|||||  
DB 420 GATGGCAGCAGCTCACTTGGAGGTACTTATGCCGCTCAGAGGCCACCTCAGAGAG 479

QY 81 LeuLeuGluGluValAlaCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
|||||  
DB 480 TTGCTTGAAGAGGTGTGTGACCGAATGAAAGAGTACGGTGAACAGATTGACCTTACC 539

QY 101 HisArgGlyAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
|||||  
DB 540 CACCGCAAGACTACGATACGCTGCGTGGAGCCGGAATGGAATCCAGTAACTAGACTTA 599

QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
|||||  
DB 600 CAGGCGATCCGAATTCACATATACAGGAGCCCTTAAGTTTCGTGTGAGAGACTT 659

QY 141 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 160  
|||||  
DB 660 GTGGAAGATACGAGAGTACGCTTATCGAATCTCTCCAGAGAGCTGAGAACGTTAA 719

QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
|||||  
DB 720 GACAAACTTGGAGTAAAGGAGACAGATCTATGTGACCAATGCCCTGACAGATCTCAGAT 779

QY 181 GluLeu 182

FEATURES

source 1. 802  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="ICRF522B2440"  
/clone\_1lb="Mouse 9-day fetus cDNA library ICRF522"  
/tissue\_type="whole embryo"  
/dev\_stage="embryonic 9-day"  
/lab\_host="E.coli, XL1 blue"  
/note="Vector: pSVSPORT1; Site\_1: NotI; Site\_2: SalI;  
Library preparation by oligo-dT priming of RNA. Clones can  
be ordered from the Resource Center in Berlin,  
http://www.rzpd.de."

BASE COUNT 193 a 206 c 243 g 159 t 1 others

ORIGIN

Tel: +49 30 8413 1612  
Fax: +49 30 8413 1380  
Email: hennig@molgen.mpg.de  
EST's are made from clones being representative of clone clusters.  
Clone clusters were calculated from oligonucleotide fingerprints.  
PCR primers  
FORWARD: 5'-GAGCATTCAGAGTACTGCA-3'  
BACKWARD: 5'-TAATACGACTCTCATAGAGG-3'  
Seq primer: 5'-ATTAGTGACATATAG-3'  
High quality sequence stop: 802.  
Location/Qualifiers

```

Db      780 GAACTT 785
|||||
RESULT 8
BI408381
LOCUS
DEFINITION
      BI408381      891 bp      mRNA      linear      EST 14-AUG-2001
      602964013F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5119563 5',
      mRNA sequence.
ACCESSION
      BI408381
VERSION
      BI408381.1  GI:15169304
KEYWORDS
      EST.
      house mouse.
SOURCE
      Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
      1 (bases 1 to 891)
      NIH-MGC http://mgs.nci.nih.gov/.
      National Institutes of Health, Mammalian Gene Collection (MGC)
      JOURNAL
      Unpublished (1999)
      Contact: Robert Strausberg, Ph.D.
      Email: cga@ds-remail.nih.gov
      Tissue Procurement: Gilbert Smith, Ph.D.
      CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
      Bernaldo, Ph.D.
      CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
      DNA Sequencing by: Incyte Genomics, Inc.
      Clone distribution: NCI-CGAP clone distribution information can be
      found through the I.M.A.G.E. Consortium/LNL at:
      http://image.llnl.gov
      Plate: L1AM1291 row: 1 column: 04
      High quality sequence stop: 873.
FEATURES
      source
          1..891
              /organism="Mus musculus"
              /strain="CZECH 11"
              /db_xref="taxon:10090"
              /clone="IMAGE:5119563"
              /clone_1lb="NCI_CGAP_Lu33"
              /tissue_type="pooled lung tumors"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: lung; Vector: pTZ19D-Pac (Pharmacia) with a
              modified polylinker; Site: 1: NotI; Site 2: EcoRI; 1st
              strand cDNA was prepared from mRNA obtained from pooled
              lung tumors with a Not I - Oligo(dT) primer [5'
              TCTTACCAATCTGAACTGGAGGGCGGCTCTCTTTTCTTTTCTTTT 3'].
              Double-stranded cDNA was ligated to Eco RI adaptors
              (Pharmacia), digested with Not I and cloned into the Not
              I and Eco RI sites of the modified pTZ19 vector. Library
              went through one round of normalization, and was
              constructed by Bento Soares and M. Fatima Bernaldo. "
BASE COUNT
      240 a      215 c      258 g      178 t
ORIGIN
Alignment Scores:
      Pred. No.:      2,76e-106      Length:      891
      Score:          946.00          Matches:      182
      Percent Similarity: 99.45%      Conservative: 0
      Best Local Similarity: 99.45%      Mismatches: 0
      Query Match:      98.64%      Indels:      1
      DB:              13      Gaps:      0
US-10-082-502-19 (1-182) x BI408381 (1-891)
OY      1      MellySGlyTTPGlyTTPLeuAlaLeuLeuLeuGlyAlaLeuLeuGlyThrAlaTTPAla 20
      |||||||
Db      198      ATGAAAGCGCGGGGTGGCGAGCCCTACTTTTGGGGGCTCTGGGAACTGCCTGGGCT 257
OY      21      ArgArgSerGlnAspLeuHISGysGlyAlaCysArgAlaLeuAlaAspGluLeuGluTTP 40
      |||||||
Db      258      CGAAGAGCCCAACATCTACCTGTGGAGCTTGCAAGGCTCTGGGTGAATTAGAGAGTGG 317
OY      41      GluITAlaArgValAspProLysTyrHISGlnMetGlySerPheArgIleAspPro 60
      |||||||

```

```

Db      318      GAAATGCCCCGCTGGACCCCAAGAACCATTTAGATGGGATCTTCGGAATCAATCA 377
OY      61      AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHISLeuThrGlu 80
      |||||||
Db      378      GATGCGAGCCAGCTCAGTTGTGGAGTACCTTATCCCGCTCAGAGGCCACTCAGAG 437
OY      81      LeuLeuGluGluValAlaCysAspArgMetGlyGlnTyrGlyGluGlnIleAspProSerThr 100
      |||||||
Db      438      TTGCTTGAAGAGGTGTGTGACCGAATGAAGAGTACGGGAAAGATTTGACCTTTCACC 497
OY      101      HISArgLysAnTYrValArgValAlaSerArgSngLysSerSerGluLeuAspLeu 120
      |||||||
Db      498      CACCGCAAGAACATCAGTACGCGGTGTGAGCCGGAATGAGAAATCAATGAATGACTTGA 557
OY      121      GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAla-CysGluSerI 140
      |||||||
Db      558      CAGGCGCATTCGAAATTGACTCGATATCAGCGCACCCCTCAAGTTTGGCTTGTGAGACAT 617
OY      140      eValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnVally 160
      |||||||
Db      618      TGTGCAAGAAATACAGGATGAGCTTATCGAATCTTCTCCAGAGAGCTGCACACGTTAA 677
OY      160      sAspLysLeuCysSerLysArgThrAspLeuCysAspHISAlaLeuHISArgSerHISAs 180
      |||||||
Db      678      AGACAAACTTGGCAGTAAAGGAGACAGATGTATGTGACCATGCTCCCTGCAGATCTCAGA 737
OY      180      pGluLeu 182
      |||||||
Db      738      TGAACCTT 744
RESULT 9
BQ210416/c      724 bp      mRNA      linear      EST 02-MAY-2002
DEFINITION
      UT-R-DY1-col-e-06-0-UT.s1 UT-R-DY1 Rattus norvegicus cDNA clone
      BQ210416
      UT-R-DY1-col-e-06-0-UT 3', mRNA sequence.
ACCESSION
      BQ210416
VERSION
      BQ210416.1  GI:20426861
KEYWORDS
      EST.
      Norway rat.
      Rattus norvegicus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
      Rattus.
      1 (bases 1 to 724)
      Bernaldo,M.F., Lennon,G. and Soares,M.B.
      Normalization and subtraction: two approaches to facilitate gene
      discovery
      Genome Res. 6 (9), 791-806 (1996)
      97044477
      Contact: Soares, MB
      Program for Rat Gene Discovery and Mapping
      University of Iowa
      451 Eckstein Medical Research Building Iowa City, IA 52242, USA
      Tel: 319 335 8250
      Fax: 319 335 9565
      Email: ms@eresblu.wieg.uiowa.edu
      The sequence contained an oligo-dT track that was present in the
      oligonucleotide that was used to prime the synthesis of first
      strand cDNA and therefore this may represent a bonafide poly A
      tail. The sequence tag present in the cDNA between the NotI site
      and the oligo-dT track served to verify it as a clone from the
      normalized cartilage library cDNA library Preparation: M.B. Soares
      Lab Clone distribution: clones will be available through Research
      Genetics (www.resgen.com)
      Seq primer: M13 Forward
      POLYA-yes.
FEATURES
      source
          1..724
              /organism="Rattus norvegicus"
              /strain="Sprague-Dawley"
              /db_xref="taxon:10116"
              /clone="UT-R-DY1-col-e-06-0-UT"
              /clone_1lb="UT-R-DY1"

```

/tissue\_type="Cartilage"  
/dev\_stage="ADULT"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Femur and Tibia; Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI; UT-R-DY1 is a normalized cDNA library  
containing the following tissue(s): Rat Cartilage from  
Femur and Tibia. The library was constructed according to  
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
1996. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was ligated to an EcoR I adaptor, digested with Not  
I, and cloned directionally into pT73-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is CTATGGACG. The Rat  
cartilage tissue was provided by Dr Jeff Stevens at the  
University of Iowa.  
TAG\_LIB-UT-R-DY1  
TAG\_TISSUE=cartilage  
TAG\_SEQ=CTATGGACG"

BASE COUNT 156 a 206 c 157 g 205 t  
ORIGIN

## Alignment Scores:

Pred. NO.: 4.79e-106 Length: 724  
Score: 943.00 Matches: 178  
Percent Similarity: 98.35% Conservative: 1  
Best Local Similarity: 97.80% Mismatches: 3  
Query Match: 98.33% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-19 (1-182) x BQ210416 (1-724)

```
Oy 1 MetLysGlyTyrPGLYTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20
Db 669 ATGAAAGGCTGGGGTGGCTGGCTTACTTTGGGGGTCTCTGGAGACTACCTGGGCT 630
Oy 21 ArgArgSerGlnAspLeuNHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTyrP 40
Db 629 CGAAGAGACGAGAGACTACACTGTGAGCTTGAGGGCTCTGGAGATGATGAGTGG 570
Oy 41 GluIleAlaAlaGlyAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 569 GAAATGGCCGGGTGGACCCCAAGAACCATTCAGATGGATCCTTCGGAATCAATCCA 510
Oy 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 509 GATGGCAGCCACTCAGTGTGTGAGAGTCTCTTATGCCCGCTCAGAGGCCACCTCAGGAG 450
Oy 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 449 TTGCTTGAGAGATATGTGACCGCATGAAAGAGTATGGGAGAACAGATTGACCTTCTACC 390
Oy 101 HisArgLysAsnTyrValAlaGlyValSerArgAsnGlyLysSerGluLeuAspLeu 120
Db 389 CACCCCAAGAACTAGCTCCGCTGTTCGCGCCGACGAGGATGCACTGAACTAGACTTA 330
Oy 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 329 CAGGGTATCCGAAATTCATTCAGATTCAGTGGCACCCTCAAGTTGCGTGTGAGGACTT 270
Oy 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160
Db 269 GTGAGAGATGACGAGAGAACTTATGATCTTTCAAGAGAGCGACAACTTAA 210
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 209 GACAAACTTTGACAGTAAACGACAGATCTATGTGACCATGCCCTCAGAGATCTCATGAT 150
Oy 181 GluLeu 182
|||||
```

Db 149 GAGCTA 144

## RESULT 10

BM632828

## LOCUS

DEFINITION

588 bp mRNA Linear EST 06-MAR-2002  
K-EST0107375 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-11-B03

ACCESSION  
BM632828  
BM632828.1 GI:19189237

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## SOURCE

## LOCATION/Qualifiers

## 1..588

## /organism="Homo sapiens"

## /db\_xref="taxon:9606"

## /clone="S5SNU484s1-11-B03"

## /clone\_lib="S5SNU484s1"

## /sex="M"

## /tissue\_type="Stomach"

## /cell\_type="Epithelial"

## /lab\_host="Top10F"

/note="Organ: Stomach; Vector: pTZ19BP1; Site 1: EcoRI;  
Site 2: NotI. The poly (A) + RNA was deapped with tabacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dT-selected  
mRNA by priming with dT-tailed vector. The dT-tailed  
vector was adjusted to have about 60nt. The cDNA vector  
was circularized with E. coli DNA ligase after digestion  
of EcoRI which site is also included in vector. An RNA  
strand converted to a DNA strand by Okayama-Berg method.  
The obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
After analyzing and sequencing about 2,000 - 3,000  
colonies in original cDNA library, the abundant cDNAs were  
selected and amplified by PCR reaction using vector region  
primer including T7 promoter as 5' primer and N(dT)14 as  
3' primer. The PCR products were used as template for  
synthesis of biotinylated single stranded RNA by in vitro  
transcription reaction. The synthesized RNA probes were  
hybridized with antisense single stranded cDNAs prepared  
from original library and incubated with avidin-gel.  
After removing DNA-RNA hybrids by centrifuge, the  
subtracted cDNA libraries were constructed by  
transformation of the remaining DNA into competent cells E.  
coli Top10F with electroporation method."

## BASE COUNT

149 a 140 c 176 g 123 t

## ORIGIN

```
Alignment Scores:
Pred. NO.: 1.47e-105 Length: 588
Score: 938.00 Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
```

Query Match: 97.81% Indels: 0  
 DB: 14 Gaps: 0  
 US-10-082-502-19 (1-182) x BM832888 (1-588)

QY 1 MetLysGlyTTPGlyTTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 2 ATGAAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 61  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspLeuLeuTrp 40  
 DB 62 CGAGAGAGCGAGAGATCTCCACTGTGGAGCATCGAGGCTCTGTGATTAACATAATG 121  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 122 GAAATTTGCCAGGTGAGCCCAAGAACACATTCAGATGGATCTTCCGATCAATCCA 181  
 QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 182 GATGCGACGCACTAGTGGTGGAGTGGCTTATGCCCTCAGAGCCCACTCCACAG 241  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
 DB 242 CTGCTGGAGGAGATATGTGACCGGATGAAGGATGTGGGAAACAGATTCCTCCACC 301  
 QY 101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyGlnSerGluLeuAspLeu 120  
 DB 302 CATGCCAAGAACTACGTACGTAGTGGCGCGGAGATGAGAACTCAGTGAACCTGACCTA 361  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 362 CAAAGCAATCCCAATCGACTCAGATATGTAGCGGACCCCAAGTTTGGCTGTGACAGCTT 421  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 DB 422 GTGAGAGAAATCGAGAGATGAATCTGATTAATCTTTTCCAGAGGCTGACAAATGTTAAA 481  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 482 GACAAACTTTCACAGTAAGCGAAGACAGATCTTGTACCATGCCCTGCACATATTCACATGAT 541  
 QY 181 GluLeu 182  
 DB 542 GAGCTA 547

RESULT 11  
 LOCUS BM832888 591 bp mRNA linear EST 06-MAR-2002  
 DEFINITION K-EST0107445 SSSNU484s1 Homo sapiens cDNA clone SSSNU484s1-11-H01  
 5' mRNA sequence.  
 ACCESSION BM832888  
 VERSION BM832888.1 GI:19189297  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 591)  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsungemall.kribb.re.kr  
 Plate: 11 row: H column: 01  
 High quality sequence stop: 591.  
 FEATURES Location/Qualifiers

source 1..591  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="SSSNU484s1-11-H01"  
 /clone\_1id="SSSNU484s1"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /lab\_host="SNU-484"  
 /note="Organ: Stomach; Vector: pTZ19mp1; site\_1: EcoRI;  
 site\_2: NotI; The poly (A) + RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 DNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okazaki-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F' by electroporation method.  
 After analyzing and sequencing about 2,000 - 3,000  
 colonies in original cDNA library, the abundant cDNAs were  
 selected and amplified by PCR reaction using vector region  
 3' primer. The PCR products were used as template for  
 synthesis of biotinylated single stranded RNA by in vitro  
 transcription reaction. The synthesized RNA probes were  
 hybridized with antisense single stranded cDNAs prepared  
 from original library and incubated with avidin-gel.  
 After removing DNA-RNA hybrids by centrifuge, the  
 subtracted cDNA libraries were constructed by  
 transformation of the remaining DNA into competent cells E.  
 coli Top10F' with electroporation method."

BASE COUNT 150 a 142 c 176 g 123 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,48e-105 Length: 591  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.358 Conservative: 2  
 Best Local Similarity: 97.258 Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 14 Gaps: 0

US-10-082-502-19 (1-182) x BM832888 (1-591)

QY 1 MetLysGlyTTPGlyTTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 2 ATGAAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 61  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspLeuLeuTrp 40  
 DB 62 CGAGAGAGCGAGAGATCTCCACTGTGGAGCATCGAGGCTCTGTGATTAACATAATG 121  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 122 GAAATTTGCCAGGTGAGCCCAAGAACACATTCAGATGGATCTTCCGATCAATCCA 181  
 QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 182 GATGCGACGCACTAGTGGTGGAGTGGCTTATGCCCTCAGAGCCCACTCCACAG 241  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
 DB 242 CTGCTGGAGGAGATATGTGACCGGATGAAGGATGTGGGAAACAGATTCCTCCACC 301  
 QY 101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyGlnSerGluLeuAspLeu 120  
 DB 302 CATGCCAAGAACTACGTACGTAGTGGCGCGGAGATGAGAACTCAGTGAACCTGACCTA 361  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140

```

|||||
Db 362 CAGGCATCCGATGACTCAGATTTAGCCGACCCCTCACTTGGCGTGAGAGCAT 421
Oy 141 VALGUGUUTYUGUASPGLUENLEUGLUPHESPESERATGUALAASPASNVALYS 160
Db 422 GTGGAGGATAGAGAGGAGTCAATGTAATCTTTCCCGAGAGCGTACAAATTTAA 481
Oy 161 ASPGLYSERGLYASRGLYTHRASPLEUCYASAPHSIALALENHISARGSERHISAP 180
Db 482 GACAAACTTTCAGTACGAGACAGATCTTTGTGACATGCCCTGCACATATCGATGAT 541
Oy 181 GUAUEN 182
Db 542 GAGCTA 547

RESULT 12
LOCUS B1712899 611 bp mRNA linear EST 11-MAR-2002
DEFINITION 1698f09.y1 HR85 1stlet Homo sapiens CDNA clone IMAGE:5085832 5'
similar to TR:Q9Y2B0 Q9Y2B0 TYPE II MEMBRANE PROTEIN.; mRNA
sequence.
ACCESSION B1712899
VERSION B1712899.1 GI:15688594
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 611)
Mellon,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemliha,I., Scarsce,M., Brestelli,J., Gradowh,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,L., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagarelshvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Research Consortium
Unpublished (2000)
CONTACT: Douglas Mellon, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmellon@idbhp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@lm.wustl.edu)
Seq primer: -40RP from GIBCO
High quality sequence stop: 463.
Location/Qualifiers
1. 611
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5085832"
/clone_lib="HR85 1stlet"
/tissue_type="Purified pancreatic 1stlet"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site:1;
NotI; Site:2; XhoI; CDNA made by oligo-dt priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."
BASE COUNT 157 a 143 c 184 g 127 t
ORIGIN
Alignment Scores: 1.55e-105 Length: 611
Pred. No.:

```

```

Score: 938.00 Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81% Indels: 0
DB: 13 Gaps: 0

US-10-082-502-19 (1-182) x B1712899 (1-611)
Oy 1 MettysgltYTPGlyTRPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20
Db 30 ATGAAGAGCGTGCGGTGGCTGGCCCTGCTTGGGGGGCCCTGCGGAACGCCCTGGCT 89
Oy 21 ArgATserGlnASPLeuHIScysGlyAlaCysATgAlaLeuValASPGLUENLUP 40
Db 90 CGGAGAGCCAGCATCTCCACTGTGAGACATGCAGCGGCTGTGGATGAACTACAAATCG 149
Oy 41 GUALEALARGVALASPProLYSLYThrILEGImetGLSerPheArgTrpLeuASPPro 60
Db 150 GAAATTCGCCAGAGTGCAGCCCAAGAAACCAATTCAGATGGATCTTCCGGATCAATCA 209
Oy 61 ASPGLYSERGLNSerValValGluValProTYRAlaARGSerGLUALAHISLeuThrGLU 80
Db 210 GATGCGACGCGAGTCACTGTGGAGTGGAGTGCCTTATGCCCGCTCAGAGGCCACCTCAGAG 269
Oy 81 LeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnILEAspProSerThr 100
Db 270 CTGCTGGAGGAGATGTGTGACCGCATGAGAGATGTGGGAAACATATGATCTTCCACCC 329
Oy 101 HISArgLYSAsnTYRValArgValValSerArgAsnGLYGLUSerSerGluLeuASPLeu 120
Db 330 CATCCCAAGAACTACGTCACGTGTACTGGCCCGAATGAGAACTCCAGTAACTGGACCTA 389
Oy 121 GlnGlyILEArgLYLeaSPSerASPLeuSerGLYThrLeuLYRPhaLAcysGluSerILE 140
Db 390 CAGGCATCCGAAATGAGTACAGATATGTAGCGGCACCTCAAGTGTGGCGTGAGAGCAT 449
Oy 141 VALGUGUUTYUGUASPGLUENLEUGLUPHESPESERATGUALAASPASNVALYS 160
Db 450 GTGGAGGATAGAGAGGATGAACATCAATCTTTTCCCGAGAGCGTACAAATTTAA 509
Oy 161 ASPGLYSERGLYSerLYArgThRASPLEUCYASAPHSIALALENHISARGSERHISAP 180
Db 510 GACAAACTTTCAGTACGAGACAGATCTTTGTGACATGCCCTGCACATATCGATGAT 569
Oy 181 GUAUEN 182
Db 570 GAGCTA 575

RESULT 13
LOCUS BF344334 697 bp mRNA linear EST 22-NOV-2000
DEFINITION 602014708P1 NCI_CGAP_Brn64 Homo sapiens CDNA clone IMAGE:4150512
5', mRNA sequence.
ACCESSION BF344334
VERSION BF344334.1 GI:11291554
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 697)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@ds-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

```







QY 121 GINGLYIleargIleasSeraspIleSerGIyThrLeuLysPheAlaCysGluSerIle 140  
DB 408 CAAGGCATCCGATGACACAGATATTTACGGCACCCTCAAGTTGGCGGTGAGACAT 467  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValys 160  
DB 468 GTGGAGGATATGAGAGATGACATGATCTTTTCCGAGAGGCTGACAAATGTTAA 527  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisArgSerHisasp 180  
DB 528 GACAAACTTGCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 587  
QY 181 GluLeu 182  
DB 588 GAGCTA 593  
RESULT 15  
LOCUS BG824549 742 bp mRNA linear EST 22-MAY-2001  
DEFINITION 602728454F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4868100 5',  
mRNA sequence.  
ACCESSION BG824549  
VERSION BG824549.1 GI:14172136  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubln Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ULNL at:  
http://image.llnl.gov  
Plate: L1CM1737 row: k column: 13  
High quality sequence stop: 740.  
Location/Qualifiers  
1. 742  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4868100"  
/clone\_1b="NIH\_MGC\_15"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GCGACAGC(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the Laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)"  
BASE COUNT 180 a 184 c 233 g 144 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,06e-105 Length: 742  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 12 Gaps: 0  
US-10-082-502-19 (1-182) x BG824549 (1-742)  
QY 1 MetLysGlyTyrGlyTyrPheAlaLeuLeuLeuGlyValLeuLeuGlyThrAlaTrrPala 20

DB 141 ATGAAGCGCTGGGGGTGGCTGGCCCTGCTTCYGGGGCCCTGGAAACCCCTGGCGCT 200  
QY 21 ArgGlySerGlyAspPheHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrP 40  
DB 201 CGAGAGCCAGCAGATCTCCACTGTGGAGCATGCAGGCTCTGTGGATGAACTTACAAATGG 260  
QY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 60  
DB 261 GAAATGGCCAGAGTGAGACCCCAAGAACCATTCAGATGGATCTTTCGGATCAATCA 320  
QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 321 GATGCAGCCAGCTCAGTGTGGAGGTGCTTATGCCCTCAGAGGCCACCTCAGAG 380  
QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr 100  
DB 381 CTGCTGGAGAGATATGTGACCCGATGAAGAGTATGGGAACAGATTGATCTTCCACC 440  
QY 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 441 CATCGCAAGACATGATGATGAGTGTAGTGGCCGGAATGGAATCCAGTGAATCGACCTA 500  
QY 121 GINGLYIleargIleasSeraspIleSerGIyThrLeuLysPheAlaCysGluSerIle 140  
DB 501 CAAGGCATCCGATGACACAGATATTTACGGCACCCTCAAGTTGGCGGTGAGACAT 560  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValys 160  
DB 561 GTGGAGGATATGAGAGATGACATGATCTTTTCCGAGAGGCTGACAAATGTTAA 620  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisArgSerHisasp 180  
DB 621 GACAAACTTGCAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 680  
QY 181 GluLeu 182  
DB 681 GAGCTA 686

Search completed: March 31, 2003, 04:51:08  
Job time : 1169.85 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 07:02:46 ; Search time 55.023 Seconds  
(without alignments)  
2614.757 Million cell updates/sec

Title: US-10-082-502-19  
Perfect score: 959  
Sequence: 1 MKGGMWLLGLLGLSTAWA.....LCSKRTDLCDAHRSNDEL 182

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 574371 segs, 425486471 residues  
Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_p2n\_model -DEV=xlp  
-Q=/cgn2.1/USPFO\_POOL/US0082502/runat\_24032003.135103\_6870/app\_query.fasta.1.1308  
-DB=Published\_Applications\_NA -QMT=fastcap -SUFFIX=trmp -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR=SCORE-pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10082502.@CGN.1.1.145 @runat\_24032003.135103.6870  
-NCP=6 -ICPU=3 -NO\_XLPEX -NO\_MMP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:

- 1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2.6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2.6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2.6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	938	97.8	814 10	US-09-864-864-309 Sequence 309, App
2	938	97.8	1210 9	US-10-028-072-173 Sequence 173, App
3	938	97.8	1210 9	US-10-121-049-173 Sequence 173, App
4	938	97.8	1210 9	US-10-123-904-173 Sequence 173, App

5	938	97.8	1210 9	US-10-140-470-173 Sequence 173, App
6	938	97.8	1210 9	US-10-175-746-173 Sequence 173, App
7	938	97.8	1210 9	US-10-176-918-173 Sequence 173, App
8	938	97.8	1210 9	US-10-176-921-173 Sequence 173, App
9	938	97.8	1210 9	US-10-137-865-173 Sequence 173, App
10	938	97.8	1210 9	US-10-140-474-173 Sequence 173, App
11	938	97.8	1210 9	US-10-142-431-173 Sequence 173, App
12	938	97.8	1210 9	US-10-143-114-173 Sequence 173, App
13	938	97.8	1210 9	US-10-140-002-173 Sequence 173, App
14	938	97.8	1210 9	US-10-142-419-173 Sequence 173, App
15	938	97.8	1210 9	US-10-123-262-173 Sequence 173, App
16	938	97.8	1210 9	US-10-142-423-173 Sequence 173, App
17	938	97.8	1210 9	US-10-121-050-173 Sequence 173, App
18	938	97.8	1210 9	US-10-141-755-173 Sequence 173, App
19	896	93.4	657 9	US-09-822-846-398 Sequence 398, App
20	262	27.3	564 10	US-09-864-864-228 Sequence 228, App
21	136.5	14.2	1457 9	US-09-864-864-228 Sequence 10, App
22	136.5	14.2	1470 9	US-10-036-041-58 Sequence 58, App
23	136.5	14.2	1470 9	US-10-035-853-58 Sequence 58, App
24	136.5	14.2	1470 9	US-10-174-590-479 Sequence 479, App
25	136.5	14.2	1470 9	US-10-176-758-479 Sequence 479, App
26	136.5	14.2	1470 9	US-10-175-737-479 Sequence 479, App
27	136.5	14.2	1470 9	US-10-173-706-479 Sequence 479, App
28	136.5	14.2	1470 9	US-10-175-728-479 Sequence 479, App
29	136.5	14.2	1470 9	US-10-175-732-479 Sequence 479, App
30	136.5	14.2	1470 9	US-10-176-482-479 Sequence 479, App
31	136.5	14.2	1470 9	US-10-176-757-479 Sequence 479, App
32	136.5	14.2	1470 9	US-10-176-913-479 Sequence 479, App
33	136.5	14.2	1470 9	US-10-180-552-479 Sequence 479, App
34	136.5	14.2	1470 9	US-10-180-557-479 Sequence 479, App
35	136.5	14.2	1470 9	US-09-931-836-58 Sequence 58, App
36	136.5	14.2	1470 9	US-10-173-700-479 Sequence 479, App
37	136.5	14.2	1470 9	US-10-174-572-479 Sequence 479, App
38	136.5	14.2	1470 9	US-10-174-574-479 Sequence 479, App
39	136.5	14.2	1470 9	US-10-174-582-479 Sequence 479, App
40	136.5	14.2	1470 9	US-10-174-588-479 Sequence 479, App
41	136.5	14.2	1470 9	US-10-175-733-479 Sequence 479, App
42	136.5	14.2	1470 9	US-10-175-744-479 Sequence 479, App
43	136.5	14.2	1470 9	US-10-175-744-479 Sequence 479, App
44	136.5	14.2	1470 9	US-10-176-488-479 Sequence 479, App
45	136.5	14.2	1470 9	US-10-176-492-479 Sequence 479, App

## ALIGNMENTS

RESULT 1  
US-09-864-864-309  
Sequence 309, Application US/09864864  
Patent No. US20020102679A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitchell, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Secrist, Heather  
APPLICANT: Lodes, Michael J.  
APPLICANT: Algate, Paul A.  
APPLICANT: Filing, Steve P.  
APPLICANT: Benson, Jane  
APPLICANT: Benson, Darin R.  
APPLICANT: Carter, Darick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.523  
CURRENT APPLICATION NUMBER: US/09/864, 864  
CURRENT FILING DATE: 2001-05-23  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: Corixa Invention Disclosure Database  
SEQ ID NO 309  
LENGTH: 814  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-864-309



;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069278  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069334  
;; PRIOR FILING DATE: 1997-12-11  
;; PRIOR APPLICATION NUMBER: 60/069694  
;; PRIOR FILING DATE: 1997-12-16  
;; PRIOR APPLICATION NUMBER: 60/072320  
;; PRIOR FILING DATE: 1998-01-23  
;; PRIOR APPLICATION NUMBER: 60/073612  
;; PRIOR FILING DATE: 1998-02-04  
;; PRIOR APPLICATION NUMBER: 60/074086  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/074092  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28

;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

Alignment Scores:  
Pred. No.: 1,01e-110 Length: 1210  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
Gaps: 0

US-10-082-502-19 (1-182) x US-10-028-072-173 (1-1210)

QY 1 MetLysGLyTrrpGLyTrrpLeuLLeuGLyValLeuGLyThrAlaTrpAla 20  
|||||  
Db 543 ATGAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602  
21 ArgArgSerGlnAspLeuHIsCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
|||||  
Db 603 CGAGAGAGCCAGGATCTCCAGTGTGAGATCGAGGCTCTGTGATACATAGATG 662  
41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
|||||  
Db 663 GAATTTGCCAGGTGAGCCCAAGAACCATTCAGATGATGATCTTCGGATCAATCCA 722  
61 AspGlySerGlnSerValValGluValProTyraIleArgSerGlnAlaHisLeuTrpGlu 80  
|||||  
Db 723 GATGCGAGCATCATGCTGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 782  
81 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 100  
|||||  
Db 783 CTGCTGAGAGAGATATGATGACCGGATGAGAGATGAGGAGAACAGATTATCTCCACC 842  
101 HisArgLysAsnTyraValArgValValSerArgAsnGlyGluSerGluLeuAspLeu 120  
|||||  
Db 843 CATCGCAAGAACTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 902  
121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuValPheAlaCysGluSerIle 140  
|||||  
Db 903 CAAGGCATCCGATGACATGACATGATATAGCGGACCCCTCAAGTTGCGTGTGAGACATT 962

OY 141 ValGluGluYrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnVallys 160  
DB 963 GTGGAGGAAATCGAGGAACTCATTAATCTTTTCCGAGAGGCTGACAAATCTTAA 1022  
OY 161 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisArgSerHisasp 180  
DB 1023 GACAACTTTCAGTACGAAACGAAAGATCTTTGTGACATGCCCTGCACATATCGCATAT 1082  
OY 181 GluLeu 182  
DB 1083 GAGCTA 1088

## RESULT 3

US-10-121-049-173

Sequence 173, Application US/10121049  
Publication No. US2003002239A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Collin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C17  
CURRENT APPLICATION NUMBER: US/10/121.049  
CURRENT FILING DATE: 2002-04-12  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-121-049-173

## Alignment Scores:

Pred. No.: 1.01e-110 Length: 1210  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-121-049-173 (1-1210)

OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGluYValLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGCTTGTGGGGCCCTGGGAAAGCGCTGGGCT 602  
OY 21 ArgArgSerGlnAspLeuHisLysGlyAlaLysArgAlaLeuValAspGluLeuGluTrp 40  
DB 603 CGGAGGAGCCGAGATCTCCATCTGTGAGCAGATGAGGCTGTGTGATGAACCTGAATGG 662  
OY 41 GluIleAlaArgValAspProLysTrpThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAATTCGCCAGGTGGAGCCCAAGAACCACTTCAGATGGGATCTTCCGGATCAATCCA 722  
OY 61 AspLysSerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 723 GATGGCAGCCAGTCAAGTGTGAGGTGCTTATGCCCGCTCAGAGGCCCACTCAGAG 782

OY 81 LeuLeuGluGluValCysAspArgMetLysGlyTrpGlyGluGlnIleAspProSerThr 100  
DB 783 CTGGTGGAGGAAATCGAGGAACTCATTAATCTTTTCCGAGAGGCTGACAAATCTTCCACC 842  
OY 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 843 CATGCCAAGAACTACGTACGTGTAGTGGCCGGAATGAGAAATCCAGTGAATCTGGACCTTA 902  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGCATCCCAATCGACTCAGATATTAAGCCGACCCCTCAAGTTGGCTGTGAGAGCATTT 962  
OY 141 ValGluGluYrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnVallys 160  
DB 963 GTGGAGGAAATCGAGGAACTCATTAATCTTTTCCGAGAGGCTGACAAATCTTAA 1022  
OY 161 AspLysLeuCySerLysArgThrAspLeuCyAspHisAlaLeuHisArgSerHisasp 180  
DB 1023 GACAACTTTCAGTACGAAACGAAAGATCTTTGTGACATGCCCTGCACATATCGCATAT 1082  
OY 181 GluLeu 182  
DB 1083 GAGCTA 1088

## RESULT 4

US-10-123-904-173

Sequence 173, Application US/10123904  
Publication No. US20030022328A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Collin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C54  
CURRENT APPLICATION NUMBER: US/10/123.904  
CURRENT FILING DATE: 2002-04-16  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-123-904-173

## Alignment Scores:

Pred. No.: 1.01e-110 Length: 1210  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-123-904-173 (1-1210)

OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGluYValLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGCTTGTGGGGCCCTGGGAAAGCGCTGGGCT 602

QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
Db 603 CGGAGGACCCAGGATCTCCACTGTGGAGCATGACAGCGCTGTGGATGAACATAAGTGG 662  
QY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 663 GAAATGGCCAGGTGGACCCCAAGAACCATTCAGATGGAGTCTTCCGGAATCATCA 722  
QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 723 GATGGACACCCAGTCAGTGTGGAGTGGCTTATGCCCCCTCAGAGCCCACTCAGACAG 782  
QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 783 CTGCTGGAGGAGATATGTGACCGGATGAGGAGATGGGAAACAGATTGATCTTCCACC 842  
QY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 843 CATCGCAAGAACTACGTACGTGTAGTGGCCCGGAATGGAAATCCAGTGAACCTGGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 903 CAAGGCAATCCGAATCGACTCAGATATTAGCGGACCCCTCAAGTTTGGCTGTGAGACGAT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerAspGluAlaAspAsnValLys 160  
Db 963 GTGAGGAAATACGAGATGAACTCATTTCTTTCCGAGAGGCTGCAATGTTAA 1022  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
Db 1023 GACAAACTTGGCAGTAAAGCAAGATCTTGTGACCATGCGCCCTGCACATATCGCATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

## RESULT 5

US-10-140-470-173  
; Sequence 173, Application US/20140470  
; Publication No. US20030022331A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C160  
; CURRENT APPLICATION NUMBER: US/10/140,470  
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-470-173

## Alignment Scores:

Pred. No.:

1.01e-110

Length:

1210

Score: 938.00  
Percent Similarity: 98.35%  
Best Local Similarity: 97.25%  
Query Match: 97.81%  
DB: 9  
Caps: 0  
Matches: 177  
Conservative: 2  
Mismatch: 3  
Indels: 0

US-10-082-502-19 (1-182) x US-10-140-470-173 (1-1210)

QY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGluValLeuGluGlyThrAlaTrpAla 20  
Db 543 ATGAAGCTGGGGTGGCTGGCCCTCTCTGCGGAGCATGACGAGGCTGTGGATGAACATAAGTGG 602  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
Db 603 CGGAGGACCCAGGATCTCCACTGTGGAGCATGACAGCGCTGTGGATGAACATAAGTGG 662  
QY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 663 GAAATGGCCAGGTGGACCCCAAGAACCATTCAGATGGAGTCTTCCGGAATCATCA 722  
QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
Db 723 GATGGACACCCAGTCAGTGTGGAGTGGCTTATGCCCCCTCAGAGCCCACTCAGACAG 782  
QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
Db 783 CTGCTGGAGGAGATATGTGACCGGATGAGGAGATGGGAAACAGATTGATCTTCCACC 842  
QY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
Db 843 CATCGCAAGAACTACGTACGTGTAGTGGCCCGGAATGGAAATCCAGTGAACCTGGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 903 CAAGGCAATCCGAATCGACTCAGATATTAGCGGACCCCTCAAGTTTGGCTGTGAGACGAT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerAspGluAlaAspAsnValLys 160  
Db 963 GTGAGGAAATACGAGATGAACTCATTTCTTTCCGAGAGGCTGCAATGTTAA 1022  
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
Db 1023 GACAAACTTGGCAGTAAAGCAAGATCTTGTGACCATGCGCCCTGCACATATCGCATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

## RESULT 6

US-10-175-746-173  
; Sequence 173, Application US/10175746  
; Publication No. US20030027270A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C353

```

: CURRENT APPLICATION NUMBER: US/10/175,746
: CURRENT FILING DATE: 2002-06-19
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-175-746-173

Alignment Scores:
Pred. No.: 1,01e-110      Length: 1210
Score: 938.00             Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81%       Indels: 0
                        Gaps: 0

US-10-082-502-19 (1-182) x US-10-175-746-173 (1-1210)

OY 1 MetLysGIYTRPGIYTRPLeuAlaLeuLeuGIYValLeuGIYThrAlaTrpAla 20
Db 543 ATGAAGCGCTGGGTGGCTGGCTGGCTGGCGGCGCTGGGAACCGCGTGGGCT 602
OY 21 ArgArgSerGIaAPLeuNIaCySGIYAlaCyArgAlaLeuValASPGLuLeuGIUtr 40
Db 603 CGGAGGACCCAGAGATCTCCAGTGGAGACAGAGGCTCTGGTGATGAATGAAAGG 662
OY 41 GluIleAlaArgValASPProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 663 GAATTGCCAGAGTGGAGCCCAAGAGACATTCAGATGGGATCTTTCCGATCAATCCA 722
OY 61 AspGlySerGlnSerValValGluValProTYrAlaArgSerGluAlaNIaNIaLeuThrglu 80
Db 723 GATGGCACCGCAGTCACTGTGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAGAG 782
OY 81 LeuLeuGIUgluValCyAspArgMetLysGIYTrgIuGluGlnIleAspProSerThr 100
Db 783 CTGCTGGAGGAGATATGTGACCCGATGAGAGATGGGAAACAGATGATGATCTTCCACC 842
OY 101 HisArgLysAsnTYrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 120
Db 843 CATGCAAGAACTACGTACGTAGTGGCGCGGAGATGAGAACTCAGTAAGTGGACCTA 902
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrIleLysPheAlaCySGluSerIle 140
Db 903 CAAGGCAATCCCAATCGATCAGATATTAGCGCACCCCAAGTTTGCCTGTGAGAGCAT 962
OY 141 ValGIUgluTYrGIUAspGIUleuIleGluPheSerArgIuAlaAspAsnValLys 160
Db 963 GTGGAGGAAATCGAGAGATGAACATTCATTTCTTTCCGAGAGGCTGACAAATGTTAAA 1022
OY 161 AspLysLeuCySerLysArgTYrAspLeuCyAspNIaLeuNIaArgSerHisAsp 180
Db 1023 GACAAACTTTCAGATTAAGCAAGATCTTTGTGACCATGCGCTGCAATATCGCATGAT 1082
OY 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 7
US-10-176-918-173
: Sequence 173, Application US/10176918
: Publication No. US20030027275A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Geritsen, Mary E.
: APPLICANT: Goddard, Audrey
```

```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P33081C382
: CURRENT APPLICATION NUMBER: US/10/176,918
: CURRENT FILING DATE: 2002-06-20
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-176-918-173

Alignment Scores:
Pred. No.: 1,01e-110      Length: 1210
Score: 938.00             Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81%       Indels: 0
                        Gaps: 0

US-10-082-502-19 (1-182) x US-10-176-918-173 (1-1210)

OY 1 MetLysGIYTRPGIYTRPLeuAlaLeuLeuGIYValLeuGIYThrAlaTrpAla 20
Db 543 ATGAAGCGCTGGGTGGCTGGCTGGCTGGCGGCGCTGGGAACCGCGTGGGCT 602
OY 21 ArgArgSerGIaAPLeuNIaCySGIYAlaCyArgAlaLeuValASPGLuLeuGIUtr 40
Db 603 CGGAGGACCCAGAGATCTCCAGTGGAGACAGAGGCTCTGGTGATGAATGAAAGG 662
OY 41 GluIleAlaArgValASPProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 663 GAATTGCCAGAGTGGAGCCCAAGAGACATTCAGATGGGATCTTTCCGATCAATCCA 722
OY 61 AspGlySerGlnSerValValGluValProTYrAlaArgSerGluAlaNIaNIaLeuThrglu 80
Db 723 GATGGCACCGCAGTCACTGTGAGTGGCTTATGCCCTCAGAGGCCACCTCAGAGAG 782
OY 81 LeuLeuGIUgluValCyAspArgMetLysGIYTrgIuGluGlnIleAspProSerThr 100
Db 783 CTGCTGGAGGAGATATGTGACCCGATGAGAGATGGGAAACAGATGATGATCTTCCACC 842
OY 101 HisArgLysAsnTYrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 120
Db 843 CATGCAAGAACTACGTACGTAGTGGCGCGGAGATGAGAACTCAGTAAGTGGACCTA 902
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrIleLysPheAlaCySGluSerIle 140
Db 903 CAAGGCAATCCCAATCGATCAGATATTAGCGCACCCCAAGTTTGCCTGTGAGAGCAT 962
OY 141 ValGIUgluTYrGIUAspGIUleuIleGluPheSerArgIuAlaAspAsnValLys 160
Db 963 GTGGAGGAAATCGAGAGATGAACATTCATTTCTTTCCGAGAGGCTGACAAATGTTAAA 1022
OY 161 AspLysLeuCySerLysArgTYrAspLeuCyAspNIaLeuNIaArgSerHisAsp 180
Db 1023 GACAAACTTTCAGATTAAGCAAGATCTTTGTGACCATGCGCTGCAATATCGCATGAT 1082
OY 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 8
```



```
US-10-176-921-173
; Sequence 173, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C28
; CURRENT APPLICATION NUMBER: US/10/176, 921
; PRIOR FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-173

Alignment Scores:
Pred. No.: 1 01e-110
Score: 938.00 Matches: 1210
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81% Indels: 0
Gaps: 0

US-10-082-502-19 (1-182) x US-10-176-921-173 (1-1210)
QY 1 Metlysglytrpqlrpleuallaleuileuileuglythraltatpala 20
Db 543 ATGAAGAGCGGGGTTGGCTGGCCCTCTCTGGGGGCGCTGCGGAACCGCTGGGCT 602
QY 21 ArgArgserlinspleuhisCysglyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 603 CGGAGGCGCAGAGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGAAGTGAATGG 662
QY 41 GluileAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 663 GAAATTCGCCAGGTGGACCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 723 GATGGCAGCCAGTCAGTGTGGAGTGCCTTATGCCCGCTCAGAGGCCCACTCAGAG 782
QY 81 LeuLeuGluLysValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
Db 783 CTCTGGAGAGATATGTACCGGATGAGAGATGAGGAGAACGATTTGATCCCTCCACC 842
QY 101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 120
Db 843 CATCGCAAGACATACGATCGTGTAGTGGCCGGAATGAGAAATCAATGAATGAGACTA 902
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGGCAATCCGAATCGACTCAGATATATAGCGGACCTCAAGTTGGGTGAGAGACTT 962
QY 141 ValGluGluTyrGluAspGluLeuIleGlnPheSerArgGluAlaAspAsnValLys 160
```

```
Db 963 GTGGAGCATACGAGATGACTCATTTTCCCGAAGGCTGACATTTTAA 1022
QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 1023 GACAACTTTCGAGTAAAGCAACAGATCTTTGTACATATCCCTGCGACATATGCAATGAT 1082
QY 181 GluLeu 182
Db 1083 GACCTA 1088

RESULT 9
US-10-137-865-173
; Sequence 173, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137, 865
; PRIOR FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-865-173

Alignment Scores:
Pred. No.: 1 01e-110
Score: 938.00 Matches: 1210
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81% Indels: 0
Gaps: 0

US-10-082-502-19 (1-182) x US-10-137-865-173 (1-1210)
QY 1 Metlysglytrpqlrpleuallaleuileuileuglythraltatpala 20
Db 543 ATGAAGAGCGGGGTTGGCTGGCCCTCTCTGGGGGCGCTGCGGAACCGCTGGGCT 602
QY 21 ArgArgserlinspleuhisCysglyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
Db 603 CGGAGGCGCAGAGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGAAGTGAATGG 662
QY 41 GluileAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 663 GAAATTCGCCAGGTGGACCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCCA 722
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 723 GATGGCAGCCAGTCAGTGTGGAGTGCCTTATGCCCGCTCAGAGGCCCACTCAGAG 782
QY 81 LeuLeuGluLysValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100
```

```
|||||
Db 783 CTGCTGGAGGAGATATGACCCGATGAGAGAGATGGGGAACACATGATCTCCACAC 842
Oy 101 HsArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 843 CATCCCAAGAACTACGACTGCTGTAGTGGCCGGAATGAGAACTCCAGTGAACCTGACCTA 902
Oy 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGCAATCCGAAATCGACTAGATATATAGCGGCAACCTCAAGTTGCGTGTAGAGCAATT 962
Oy 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 963 GTGGAGGAATACGAGATGAACTCATTAATCTTTCCCGAGAGGCTGACAAATGTTAAA 1022
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 1023 GACAAACTTTGCAGTAACGAAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
Oy 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 10
US-10-140-474-173
; Sequence 173, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-173

Alignment Scores:
Pred. No.: 1,01e-110 Length: 1210
Score: 938.00 Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-140-474-173 (1-1210)
Oy 1 MetLysGlyTyrGlyTyrPheAlaLeuLeuLeuGlyValLeuLeuGlyThrAlaTyrPala 20
Db 543 ATGAAGAGCTGGGCTGGCTGGCTGTGGGGGCGCTGCTGGAGAACCGCGTGGGCT 602
Oy 21 ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValAspGluLeuGluTyrP 40
```

```
|||||
Db 603 CGAGAGACCCAGGACTCTCACACTGTGAGACATGACAGGCTCTGTGGATGAACTAGAAATGG 662
Oy 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60
Db 663 GAAATTTGCCAGGTGTGACCCCAAGAAACATTTTCAGATGGAGATCTTCCGGATCAATCCA 722
Oy 61 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80
Db 723 GATGCGACCGAGTCAAGTGTGTGAGGTGCTTATGCCCCTCGAGAGGCCACCTGCACAGAG 782
Oy 81 LeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100
Db 783 CTGCTGGAGGAGATATGACCCGATGAGAGAGATGGGGAACACATGATCTCCACAC 842
Oy 101 HsArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 843 CATCCCAAGAACTACGACTGCTGTAGTGGCCGGAATGAGAACTCCAGTGAACCTGACCTA 902
Oy 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 903 CAAGCAATCCGAAATCGACTAGATATATAGCGGCAACCTCAAGTTGCGTGTAGAGCAATT 962
Oy 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160
Db 963 GTGGAGGAATACGAGATGAACTCATTAATCTTTCCCGAGAGGCTGACAAATGTTAAA 1022
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 1023 GACAAACTTTGCAGTAACGAAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082
Oy 181 GluLeu 182
Db 1083 GAGCTA 1088

RESULT 11
US-10-142-431-173
; Sequence 173, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173

Alignment Scores:
Pred. No.: 1,01e-110 Length: 1210
Score: 938.00 Matches: 177
```

Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-142-431-173 (1-1210)

QY 1 MetlysglytrpplytrpleuallaleuLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAAGCGTGGGTTGGCTGGCCCTCTCTGGGGCCCTCGGGAACCGCTGGGCT 602

QY 21 ArgArSerglnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCAAGGCTGTGGTGATGAAGTGG 662

QY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAATTGCCAGGTGAGCCCAAGAGACCATTCAGATGGAGATCTTCGGGATCAATCCA 722

QY 61 AspGlySerGlnSerValValGluValProTyraAlaArgSergluAlaHisLeuThrGlu 80  
DB 723 GATGGCAGCCAGTCAGTGGTGGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAG 782

QY 81 LeuLeuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSertThr 100  
DB 783 CTGCTGGAGAGATATGTGACCGGATGAGAGATGAGGAGATGATCCTTCACAC 842

QY 101 HisArgLysAsnTyraValaArgValValSerArgAsnGlyLysSerSergluLeuAspLeu 120  
DB 843 CACGCGAAGAACTACGTACGTAGTGGCCGGAATGGAATCCAGTGAACAGGACCTA 902

QY 121 GluGlyIleArgLysAspSerAspIleSerglyThrLeuLysPheAlaCysGluSertThr 140  
DB 903 CAAGGCGTCCGAATCGCATGATATTAGCGGACCTCAAGTTGGCTGTGAGACATT 962

QY 141 ValGluGluTrpGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 963 GTGAGGAATACGAGATGATGACATTTCTTTCCGAGAGGCTGCAATGTTAA 1022

QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSertHisAsp 180  
DB 1023 GACAACTTGGCAGTAGGACGACAGATCTTTGTGACCATCCCTGCACATATGCATGAT 1082

QY 181 GluLeu 182  
DB 1083 GAGCTA 1088

RESULT 12  
US-10-143-114-173  
Sequence 173, Application US/10143114  
Publication No. US20030036180A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria A.  
APPLICANT: Stewart, Daniel  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C211  
CURRENT APPLICATION NUMBER: US/10/143,114

CURRENT FILING DATE: 2002-05-09  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-143-114-173

Alignment Scores:  
Pred. No.: 1,016-110 Length: 1210  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-143-114-173 (1-1210)

QY 1 MetlysglytrpplytrpleuallaleuLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
DB 543 ATGAAAGCGTGGGTTGGCTGGCCCTCTCTGGGGCCCTCGGGAACCGCTGGGCT 602

QY 21 ArgArSerglnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
DB 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCAAGGCTGTGGTGATGAAGTGG 662

QY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAATTGCCAGGTGAGCCCAAGAGACCATTCAGATGGAGATCTTCGGGATCAATCCA 722

QY 61 AspGlySerGlnSerValValGluValProTyraAlaArgSergluAlaHisLeuThrGlu 80  
DB 723 GATGGCAGCCAGTCAGTGGTGGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAG 782

QY 81 LeuLeuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSertThr 100  
DB 783 CTGCTGGAGAGATATGTGACCGGATGAGAGATGAGGAGACAGTGTATCCTTCACAC 842

QY 101 HisArgLysAsnTyraValaArgValValSerArgAsnGlyLysSerSergluLeuAspLeu 120  
DB 843 CACGCGAAGAACTACGTACGTAGTGGCCGGAATGGAATCCAGTGAACAGGACCTA 902

QY 121 GluGlyIleArgLysAspSerAspIleSerglyThrLeuLysPheAlaCysGluSertThr 140  
DB 903 CAAGCAGTCCGAATCGCATGATATTAGCGGACCTCAAGTTGGCTGTGAGACATT 962

QY 141 ValGluGluTrpGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 963 GTGAGGAATACGAGATGATGACATTTCTTTCCGAGAGGCTGCAATGTTAA 1022

QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSertHisAsp 180  
DB 1023 GACAACTTGGCAGTAGGACGACAGATCTTTGTGACCATCCCTGCACATATGCATGAT 1082

QY 181 GluLeu 182  
DB 1083 GAGCTA 1088

RESULT 13  
US-10-140-002-173  
Sequence 173, Application US/10140002  
Publication No. US20030037623A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P330R1C59  
CURRENT APPLICATION NUMBER: US/10/140,002  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-002-173

Alignment Scores:  
Pred. No.: 1,01e-110 Length: 1210  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: Gaps: 0

US-10-082-502-19 (1-182) x US-10-140-002-173 (1-1210)

OY 1 MetLysGlyTyrPGLYTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrrPa 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGGCTGGGGGCGCCCTGGGAACCGCTGGGCT 602  
OY 21 ATGATGSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrP 40  
DB 603 CGAGAGACCGAGATCTCCACGTGAGACATGCAGGCTCTGGTGAGAACTAGAAATGG 662  
OY 41 GlnLeaAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAATTTGCCAGAGTGGAGCCCAAGAAAGACATTCAGATGGAGTCTTCGGATCAATCA 722  
OY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHiscLeuThrGlu 80  
DB 723 GATGGACCCAGATGATGTGACCGGATGAAGAGATGGGGAACAGATTCCTTCACAC 842  
OY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 100  
DB 783 CTGCTGGAGAGATATGTGACCGGATGAAGAGATGGGGAACAGATTCCTTCACAC 842  
OY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 843 CATCCAGAACTACGTACGTACTAGTGGCGGAAATGGAAATCCAGTAAGTACCTA 902  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGGACCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
DB 963 GTGGAGAAATAGAGAGATGAACTCTTAATTTCTTTCCGAGAGGCTGACATGTTAA 1022  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
DB 1023 GACAAACTTTCAGTAAGCAAGACAGATCTTTGTGACATGCGCTGCATATCGCATAT 1082  
OY 181 GluLeu 182  
DB 1083 GAGCTA 1088  
RESULT 14  
US-10-142-419-173

Sequence 173, Application US/10142419  
Publication No. US20030044945A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Mei-Qiang  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P330R1C244  
CURRENT APPLICATION NUMBER: US/10/142,419  
CURRENT FILING DATE: 2002-05-10  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 173  
LENGTH: 1210  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-142-419-173

Alignment Scores:  
Pred. No.: 1,01e-110 Length: 1210  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: Gaps: 0

US-10-082-502-19 (1-182) x US-10-142-419-173 (1-1210)

OY 1 MetLysGlyTyrPGLYTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrrPa 20  
DB 543 ATGAAGGCTGGGGTGGCTGGCTGGCTGGGGGCGCCCTGGGAACCGCTGGGCT 602  
OY 21 ATGATGSerGlnAspLeuHiscysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrP 40  
DB 603 CGAGAGACCGAGATCTCCACGTGAGACATGCAGGCTCTGGTGAGAACTAGAAATGG 662  
OY 41 GlnLeaAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 663 GAATTTGCCAGAGTGGAGCCCAAGAAAGACATTCAGATGGAGTCTTCGGATCAATCA 722  
OY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHiscLeuThrGlu 80  
DB 723 GATGGACCCAGATGATGTGACCGGATGAAGAGATGGGGAACAGATTCCTTCACAC 842  
OY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 100  
DB 783 CTGCTGGAGAGATATGTGACCGGATGAAGAGATGGGGAACAGATTCCTTCACAC 842  
OY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
DB 843 CATCCAGAACTACGTACGTACTAGTGGCGGAAATGGAAATCCAGTAAGTACCTA 902  
OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 903 CAAGGACCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962  
OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160

Db 963 GTGAGAGATACGAGATGACATCATGATCTTTCCCGAGAGCTGACATGTTAA 1022  
QY 161 AsplysLeuCySerIysArgThrAspLeuCyAspHisAlaLeuHisArgSerHisAsp 180  
Db 1023 GACAACTTTGACAGTACGAGACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

## RESULT 15

US-10-123-262-173  
: Sequence 173, Application US/10123262  
: Publication No. US20030049816A1  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin P.  
: APPLICANT: Beresini, Maureen  
: APPLICANT: DeForge, Laura  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Gao, Wei-Qiang  
: APPLICANT: Gerltzen, Mary E.  
: APPLICANT: Goddard, Audrey  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Sherwood, Steven  
: APPLICANT: Smith, Victoria  
: APPLICANT: Stewart, Timothy A.  
: APPLICANT: Tumas, Daniel  
: APPLICANT: Watanabe, Colin K  
: APPLICANT: Wood, William  
: APPLICANT: Zhang, Zemin  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
: FILE REFERENCE: P330R1C38  
: CURRENT APPLICATION NUMBER: US/10/123,262  
: CURRENT FILING DATE: 2002-04-15  
: Prior Application removed - See File Wrapper or Palm  
: NUMBER OF SEQ ID NOS: 550  
: SEQ ID NO 173  
: LENGTH: 1210  
: TYPE: DNA  
: ORGANISM: Homo Sapien  
US-10-123-262-173

## Alignment Scores:

Pred. No.: 1.01e-110 Length: 1210  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x US-10-123-262-173 (1-1210)

QY 1 MetIysGLYTPGLYTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTTPAla 20  
Db 543 ATGAAGGCTGGGGTGGTGGCCCTCTCTGGGGCCCTGCTGGAACCCCTGGGCT 602  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTTP 40  
Db 603 CGAGAGGCCAGGATCTCCACTGTGAGCAAGGAGGCTGCTGGTGAATGAGATG 662  
QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
Db 663 GAAATGGCCAGGTGACCCCAAGAACCAATCAGATGGATCTTCCGATCAATCCA 722  
QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
Db 723 GATGGCAGCCAGTCAAGTGTGAGAGTGCTTATGCCGCTCAGAGGCCACCTCAGAG 782  
QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100

Db 783 CTGCTGAGAGATATGTCACCGGATGAAGAGATATGGGCAACAGATTGATCTTCCACC 842  
QY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
Db 843 CATCGCAAGAACTACGTACGTGTAGTGGCCGCGATGAGAAATCCAGTAACCTGACCTA 902  
QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
Db 903 CAAGGCATCCGAAATCGACTCAGATTTTACCGGCAACCTCAAGTTTGGCTGTGAGAGCAT 962  
QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
Db 963 GTGAGAGATACGAGATGAACTCATGATTTCTTTCCCGAGAGGCTGACAAATGTTAA 1022  
QY 161 AsplysLeuCySerIysArgThrAspLeuCyAspHisAlaLeuHisArgSerHisAsp 180  
Db 1023 GACAACTTTGACAGTACGAGACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 1082  
QY 181 GluLeu 182  
Db 1083 GAGCTA 1088

Search completed: March 31, 2003, 11:59:08  
Job time: 59.0233 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 39.6802 Seconds

(without alignments)  
1406.624 Million cells updates/sec

Title: US-10-082-502-19

Sequence: 1 MKGCMGLALLGLVLTGTAWA.....LCSKRTDLCDAHNSHDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL-frame\_p2n.model -DEV-xlp  
-Q-/cgn2\_1/USPTO/US10082502/runat\_24032003.135100.6616/app\_query.fasta\_1.1308  
-DB-Issued\_Patents\_NA -QFMT-fastap -SUFFIX-rn1 -MINMATCH-0.1 -LOOPCL-0  
-LOOPEXT-0 -UNITS-Dits -SPART-1 -END-1 -MATRIX-Dlosum62 -TRANS-human40.cdl  
-LIST-45 -DOCALLIGN-200 -THR SCORE-pct -THR MAX-100 -THR MIN-0 -ALIGN-15  
-NODE-LOCAL -OUTFMT-pio -NOR-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000  
-USER-US10082502-ecgn1\_1.131-ecgnat\_24032003.135100.6616 -NCPU-6 -ICPU-3  
-NO\_XLPUX -NO\_MAP -LARGEQUERY -NEG\_SCORES-0 -WAIT -LONGLOP -DEV\_TIMEOUT-120  
-NARN\_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7  
-XGAPOP-10 -XGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database: Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/6C.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/6D.COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136.5	14.2	744	4	US-09-163-285-3
2	136.5	14.2	1512	4	US-09-163-285-1
3	100	10.4	915	4	US-09-724-864-14
4	87	9.1	5181	1	US-08-257-073-10
5	84	8.8	1530	4	US-09-131-831B-2
6	84	8.8	1554	3	US-08-809-999D-1
7	84	8.8	1554	4	US-09-069-637-1
8	84	8.8	1554	4	US-09-322-360-1
9	84	8.8	1554	4	US-09-131-831B-1
10	84	8.8	1555	2	US-08-316-231B-1
11	84	8.8	2148	3	US-08-809-999D-2
12	84	8.8	2148	3	US-09-069-637-2

13	84	8.8	2148	4	US-09-322-360-2	Sequence 2, Appl1
14	82.5	8.6	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
15	82.5	8.6	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
16	82	8.6	6519	1	US-08-588-985-1	Sequence 1, Appl1
17	82	8.6	6519	1	US-08-971-988-1	Sequence 62, Appl1
18	81.5	8.5	61663	4	US-09-453-702B-62	Sequence 38, Appl1
19	79.5	8.3	5661	4	US-08-938-105-2	Sequence 38, Appl1
20	78	8.1	4131	1	US-07-872-644-38	Sequence 38, Appl1
21	78	8.1	4131	1	US-08-297-494-38	Sequence 38, Appl1
22	78	8.1	4131	1	US-08-297-510-38	Sequence 38, Appl1
23	78	8.1	4131	1	US-08-479-532-38	Sequence 38, Appl1
24	78	8.1	4131	1	US-08-455-526-38	Sequence 38, Appl1
25	78	8.1	4131	1	US-08-455-526-38	Sequence 38, Appl1
26	78	8.1	4131	1	US-09-139-491-38	Sequence 38, Appl1
27	78	8.1	4131	5	PCT-US92-03222-38	Sequence 38, Appl1
28	77.5	8.1	1924	4	US-08-961-083-159	Sequence 159, App
29	76.5	8.0	3752	4	US-08-961-527-208	Sequence 208, App
30	76	7.9	833	2	US-08-837-029-1	Sequence 1, Appl1
31	76	7.9	3789	1	US-07-872-644-42	Sequence 42, Appl1
32	76	7.9	3789	1	US-08-297-494-42	Sequence 42, Appl1
33	76	7.9	3789	1	US-08-297-510-42	Sequence 42, Appl1
34	76	7.9	3789	1	US-08-479-532-42	Sequence 42, Appl1
35	76	7.9	3789	1	US-08-455-526-42	Sequence 42, Appl1
36	76	7.9	3789	1	US-08-455-526-42	Sequence 42, Appl1
37	76	7.9	3789	3	US-09-139-491-42	Sequence 42, Appl1
38	76	7.9	3789	5	PCT-US92-03222-42	Sequence 103, App
39	76	7.9	5558	4	US-08-961-527-103	Sequence 3, Appl1
40	75.5	7.9	1216	4	US-09-071-035-3	Sequence 1, Appl1
41	75.5	7.9	1347	4	US-09-071-035-1	Sequence 1, Appl1
42	74.5	7.8	4371	1	US-08-803-972-1	Sequence 1, Appl1
43	74.5	7.8	4371	1	US-08-803-972-1	Sequence 9, Appl1
44	74	7.7	918	3	US-08-937-271-9	Sequence 20, Appl1
45	74	7.7	1349	1	US-07-951-715A-20	

## ALIGNMENTS

RESULT 1  
US-09-163-285-3  
; Sequence 3, Application US/09163285  
; Patent No. 6204013  
; GENERAL INFORMATION:  
; APPLICANT: Rhododoust, Mehran  
; TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/163,285  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/090,398  
; FILING DATE: June 24, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragoras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 3:



```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 744 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..744
: US-09-163-285-3
:
: Alignment Scores:
: Pred. No.: 6,43e-08 Length: 744
: Score: 136.50 Matches: 49
: Percent Similarity: 44.68% Conservative: 35
: Best local Similarity: 26.06% Mismatches: 61
: Query Match: 14.23% Indels: 43
: DB: Gaps: 10
:
: US-10-082-502-19 (1-182) x US-09-163-285-3 (1-744)
:
: Oy 11 LeuGIuValLeuLeu-----GlyThrAlaTrpAla-----Arg 21
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 16 TTGGGAATATTGCTCTTCCCTTTTGGCCCTGCACAGAGCTTGCGGTGGATCTTGAAG 75
:
: Oy 22 ArgSerGlnAsp-----LeuHisCysGlyAlaCysArgAlaLeuVal 35
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 76 GAGGAGCAGCTGTGACACAGAACGCTTCCCGCAAAATGCCAAGGTGTAACTGCTGACC 135
:
: Oy 36 AspGluLeuGlnGluTrpGluIleAlaArgValAsp---ProTyrSlyTrpIleGlnMetGly 54
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 136 ACAGAGCTACACGGCGGAGAACTAGTCCACCGCTGATCTCGAGAGGTCTGGACCTGGG 195
:
: Oy 55 SerPheArgIleAsnProAspGlySerGlnSerValValGluValProTyrAlaArgSer 74
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 196 CAGGTGCTG-----GATACAGCGCAAGAGAAAGACACAGCTGCTTACAGCGTTTCA 246
:
: Oy 75 GluAlaHisLeuThrGluLeuLeuGluGluValCysAspTrpMetLysGluTyrGlyGlu 94
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 247 GAGACAAAGGCTGGAAGAGCGCTTATAGAGATTATGTAGCGCGAGCTCGACTAT----- 300
:
: Oy 95 GlnIleAspProSerThrHis-----ArgLysAsnTyrValAlaArgValIleSerArgAsn 112
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 301 -----ACTGTTTACAGCTGAGCGCAAGGGCTCAGTCAAGATATTCGCAAGGCTCAG 348
:
: Oy 113 GlyGluSer-----SerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 349 AGTAGAGACATGGCACAACCTGAAGAGCGCTTACTGACAGAGGGGGGGAAGGTGATCTGGG 408
:
: Oy 129 IleSer-----GlyThrLeuLysPheAlaCys 137
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 409 ATCCCTCTGGAGAGCTTTGGGATGAGACCCAGCGTGGAGGTCACTACCTCCAAAGACATGT 468
:
: Oy 138 GluSerIleValIleGluGluTyrGluAspGluLeuIleGlu---PhePheSerArgGluAla 156
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 469 GAGACCAAGTGTGGAGAGATTGTGAAGACATGTGGGAGACATGTGTTCTTCCACATCAGAG 528
:
: Oy 157 AspaAsnValLysAspLysLeuCys 164
: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 529 CAGCCCTTACAAATTTTCTCTGT 552
:
: RESULT 2
: US-09-163-285-1
: Sequence 1, Application US/09163285
: Patent No. 6204013
: GENERAL INFORMATION:
: APPLICANT: Khodadoust, Mehran
: TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES
: TITLE OF INVENTION: AND USES THEREOF
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHYE & COCKFIELD, LLP
: STREET: 28 State Street
:

```

```

      CITY: Boston
      STATE: Massachusetts
      COUNTRY: USA
      ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM type: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/163,285
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/090,398
FILING DATE: June 24, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: MNI-049
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1512 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 125..868
US-09-163-285-1

Alignment Scores:
Pred. No.: 1,79e-07 length: 1512
Score: 136.50 Matches: 49
Percent Similarity: 44.68% Conservative: 35
Best Local Similarity: 26.06% Mismatches: 61
Query Match: 14.23% Indels: 43
DB: 4 Gaps: 10
US-10-082-502-19 (1-182) x US-09-163-285-1 (1-1512)

QY 11 LeuGlyValLeuLeu-----GlyThrAlaTrpAla-----Arg 21
   |||||:|||||
Db 140 TTGGGAATATTGCTTTTCCTTTTGTGGCCGTGCACGAGCGTGGCGGAGTGTGCAAG 199
   |||||:|||||

QY 22 ArgSerGlnAsp-----LeuHisCysGlyAlaCysArgAlaLeuVal 35
   |||||:|||||
Db 200 GAGGAGGACGATGACACAGAACGCTTGCCCAAGCAATGCCAAGTGTTAAGCTGTAGC 258
   |||||:|||||

QY 36 AspGluLeuGluTrpGluIleAlaArgValAsp--ProLysIuYrThrIleGlnMetGly 54
   |||||:|||||
Db 260 ACAGAGCTACAGCGCGGAACCTAGTCGACCGCTGCATCTCGAGAGGAGTGTGAGCTGGG 319
   |||||:|||||

QY 55 SerPheArgIleAsnProAspGlySerGlnSerValValGluValProTyrAlaAspSer 74
   |||||:|||||
Db 320 CAGGTGCTG-----GATPACAGGCAAGAGAGACACGTCCTTACAGCGTTTCA 370
   |||||:|||||

QY 75 GluAlaHisLeuThrGluLeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlu 94
   |||||:|||||
Db 371 GAGACAAAGCGCTGGGAAGAGCGCTTATGAGAAATTATGATGAGGAGATCCCGACTAT----- 424
   |||||:|||||

QY 95 GlnIleAspProSerThrHis-----ArgLysAsnTyrValArgValValSerArgAsn 112
   |||||:|||||
Db 425 -----AGTGTTCACAGCTGTGAGCGCAAGGCGCTCATCTGATGATATGCCAAGGTCAG 472
   |||||:|||||

QY 113 GlyGluSer-----SerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128
   |||||:|||||
Db 473 AGTCAGACCATGGCAACACTGAAAGCCCTAGTCGACGAAGGGGGGTAAGGTGATCTGGGG 532
   |||||:|||||

```



US-10-082-502-19 (1-182) x US-08-257-073-10 (1-5181)

Oy 36 AspGluLeuGluTyrPgiuIleAlaArgValAspProLysrLysThrIleGlnMetGlySer 55  
||| ||||||||| ||| : : : ||| : : :  
Db 1192 GATCACCCTGGATTGAATTATTTAAGAGAAAAAAAAGTTGTATACACT 1251

Oy 56 PheArgIleAsnProAspGlySerGlnSerValValGluValProTyraIlaArgSergLu 75  
: : : ||| : : : ||| : : :  
Db 1252 AAATCAACAAGATCCTAGCAAAATCGTTCAATACCAAATACTTCCTTAATCGATTT 1311

Oy 76 AlaHis-----LeuthrGluLeuLeuGluGluValcys-----AspargmetLys 90  
: : : ||||||: : : : : : :  
Db 1312 GTATATCCCTTACACACTGCATCATTAATTCATTAATCTGCAGATATGATTAATAAT 1371

Oy 91 GluTyrGluGluGlnIleAspProSerThiHlaArgLysAsnTyraIaArgValValSer 110  
||| | : : : ||| : : :  
Db 1372 TCATATGGTGATTAAAGAATCCGATACTAAGAAAAAATTAATGAATAAATATTATCA 1431

Oy 111 ArgAsnGluGlu-----SerSergLeuAspLeuGlnGly 122  
||| || : : : ||| : : :  
Db 1432 GATTAATAAGAAAGAAAAAATTCATTATTAACATTAATAAACAAATGATTTAGAAAGA 1491

Oy 123 IleArgIleAspSerAspIleSergLythrLeuLysPheAlaCysGluSerIleValGlu 142  
: : : ||| : : : ||| : : :  
Db 1492 AAAAAAATTATATCAC-----ACAAAGAACMAAAATATAAATAATTAATCTGAA 1536

Oy 143 GluTyrGlu-----AspGluLeuIleGlu---PhePheSerArgGlu 155  
: : : ||| : : : ||| : : :  
Db 1537 GATTATGAAGAAAGTCAAAAAAGATTATGAAGAAATCTTGAAAAATTTTAGAAATGAA 1596

Oy 156 AlaAspAsn-----ValLysAspLysLeuCySerLysArg 167  
: : : ||| : : : ||| : : :  
Db 1597 TTTAATATATATTTTGACAAAGATGCTGATATTAATAATTCAGTGCAGAA 1647

RESULT 5  
US-09-131-831B-2  
; Sequence 2, Application US/09131831B  
; Patent No. 6339149  
GENERAL INFORMATION:  
APPLICANT: Coulie, Pierre; Ikeda, Hideyuki; Boon-  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
Coding For Tumor Rejection Antigen Precursors DAGE and  
Uses Thereof  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/131.831B  
FILING DATE: 11-Aug-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,999  
FILING DATE: 9-April-1997  
APPLICATION NUMBER: 08/316,221  
FILING DATE: 30-September-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, NO. 6339149man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3100

```

TELEFAX: (212) 318-3400
:
: INFORMATION FOR SEQ ID NO: 2:
:
: SEQUENCE CHARACTERISTICS:
:
:   LENGTH: 1530 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: double
:   TOPOLOGY: linear
:
: MOLECULE TYPE: nucleic acid
:
: FEATURE:
:
:   NAME/KEY:   DATE (H12)
:   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
:
US-09-131-831B-2

Alignment Scores:
Pred. No.:      0.624      Length:      1530
Score:          84.00      Matches:      40
Percent Similarity: 38.67%  Conservative: 30
Best Local Similarity: 22.10% Mismatches:      61
Query Match:      8.76%     Indels:       50
DB:              4         Gaps:         8

US-10-082-502-19 (1-182) x US-09-131-831B-2 (1-1530)
Oy      1  MetLysGIYTPrgLYTPrLeuAlaLeuLeuLeuGIYValLeuLeuGIYThrAlaTrpAla 20
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      217  GTGAGGCGCTGGCCCTCTACCTGCTCCCTCTGGAGTGGCTGATG----- 261
Oy      21  ArgArgSerGIaAspLeuHisCysGIYAlaCysArgAlaLeuValaAspGIuLeuGIu--- 39
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      262  ---AAGGACAAACATCTTCACCTGGAGACCTTCAAAGCTGTGCTGATGACCTGATGTG 318
Oy      40  -----TrpIuIleAlaArgValaAspProLys 49
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      319  CTCCTTGCCACGAGAGTTCGCCCCAGAGGTGGAAACTTCAGTGTGATTAACGGAG 378
Oy      50  -----ThrIleGluMetGIY-----SerPhe 56
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      379  AACCTCTATCAGACTCTTGACTGTGTATGTCTGTGAACAGGCCACTCTGTACTATT 438
Oy      57  -----ArgIleAsnProAspGIYSerGlnSer 65
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      439  CCAGAGCCAGAAAGCAGCTCAGCCCATGACAAAGAAAGGAAAGTATGATGGTTGACACA 498
Oy      66  ValValGIuValaProTYrAlaArgSerGIuAlaHisLeuThrGIuLeuLeuGIu--- 84
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      499  GAGCGACAGAGCCCTTCATTCACAGTAGAGGTGCTCTCAGCCCTGTTCCTCAAGGAAGGT 558
Oy      85  ValCysAspArgMetLysGIuTYrGIYGIuGlnIleAspProSerThrHisArgLysAsn 104
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      559  GCCTGTGATGATGATTGTTCTCTAC-----CTCATTAAGAAAGTGAAGCAAGAAATAAT 612
Oy      105  TYrValArgValaIleSerArgAsnGIYGIuSerSerGIuLeuValaAspLeuGlnGIYIleArg 124
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      613  GTACTAGCCCGCTGCTGTAAGAAAGCTGAAGATTTTTCATAGCCCAAGCAG----- 663
Oy      125  IleAspSerAspIleSerGIYThrLeuLysPheAlaCysGIuSerIleValaGIuGIuTYr 144
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      664  -----GATATCAAGATGATGATCTGAATAATGCTGCAGCTGCAGCTCATTTGAAGATTG 714
Oy      145  Glu 145
      :  :  :
Db      715  GAA 717

RESULT 6
US-08-809-999D-1
: Sequence 1, Application US/08809999D
: Patent No. 6013765
:
: GENERAL INFORMATION:
:
: APPLICANT: Coulie, Pierre; Ikeda, Hideyuki;
: APPLICANT: Boon-Fallou, Thierry
:
: TITLE OF INVENTION: Isolated Nucleic Acid Molecules
:
: TITLE OF INVENTION: Coding for Tumor Rejection Antigen Precursors
:
: DATE OF INVENTION: Uses thereof

```

NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski, L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,999D  
FILING DATE: 9-April-1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/316,231  
FILING DATE: 30-September-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6013765man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE (5E10)  
US-08-809-999D-1

Alignment Scores:  
Pred. No.: 0.638 Length: 1554  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61  
Query Match: 8.76% Indels: 50  
DB: 3 Gaps: 8

US-10-082-502-19 (1-182) x US-08-809-999D-1 (1-1554)

QY 1 MetLysGLYTrpGLYTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
DB 276 GTGCAAGGCTGGCCCTTCACTGCTCCCTCTGGAGAGTGTGATG----- 320  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlu--- 39  
DB 321 ---AAGGACACATCTTCACCTGAGACCTTCAAAAGCTGTGTTGATGAGCTGATGTC 377  
QY 40 -----TrrGluIleAlaArgValAspProLysLys 49  
DB 378 CTCCTTGGCCAGAGGTTGCCCCAGAGAGGTGGAACCTTCAAGTCTGATTTACGGAAG 437  
QY 50 -----ThrlleGlnMetGly-----SerIle 56  
DB 438 AACTCTCATCAGAGACTTCTGAGCTGTATGCTGTGGAAACAGGCCAGCTGTACTCATTT 497  
QY 57 -----ArgIleAsnProAspGlySerGlnSer 65  
DB 498 CCAGAGCCAGAGACACTCAGCCCATGACAAAGAGCGAAAGATGATGTTTGACACACA 557  
QY 66 ValValGluValProGlyTrpAlaArgSerGluAlaHisLeuThrGluLeuGluGlu--- 84  
DB 558 GAGGACAGACACCCCTTCATTCAGTAGAGAGTGTCTGTAGACCTGTCTTCAAGGAAGGT 617  
QY 85 ValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThrHisArgLysAsn 104

DB 618 GCCTGATGATAATGTTCTCTCTAC-----CTCATTTAGAAAGTGAAGCCAGAAAGAAAAAT 671  
QY 105 TyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArg 124  
DB 672 GTACTAGCGCTGTGCTGTAAAGACCTGAAGATTTTTCATTCACCAATGCCACGAG----- 722  
QY 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTrp 144  
DB 723 -----GATATCAAGATGATCTCTGAAATAGGTGACGTGAGACTGTATTAAGATTTG 773  
QY 145 Glu 145  
DB 774 GAA 776

RESULT 7  
US-09-069-637-1  
Sequence 1, Application US/09069637  
Patent No. 6022692  
GENERAL INFORMATION:  
APPLICANT: Coullie, Pierre; Ikeda, Hideyuki;  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
TITLE OF INVENTION: Coding For Tumor Rejection Antigen Precursors DAGE and Uses  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/069,637  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,999  
FILING DATE: 9-April-1997  
APPLICATION NUMBER: 08/316,231  
FILING DATE: 30-September-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6022692man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE (5E10)  
US-09-069-637-1

Alignment Scores:  
Pred. No.: 0.638 Length: 1554  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61  
Query Match: 8.76% Indels: 50  
DB: 3 Gaps: 8

US-10-082-502-19 (1-182) x US-09-069-637-1 (1-1554)

Oy 1 MetLysGlyTyrPheLysIleuAlaLeuLeuGlyValLeuLeuGlyThrAlaTyrPhe 20  
 Db 276 GTGGAGGCGCTGGCCCTTCACCTGCTCCCTGGAGTGGCTGATG----- 320  
 Oy 21 ArgArgSerGlnAspLeuHisCysGlyValAlaCysArgAlaLeuValAspGluLeuGlu--- 39  
 Db 321 ---AAGGAGCAACAATCTTCACACCTGGAACTTCAAAGCTGCTGCTGATGGACCTGATG 377  
 Oy 40 -----TyrGluIleAlaArgValAspProLys 49  
 Db 378 CTCCTTGCCAGAGGTCGCCCCAGAGAGGTGGAACTTCAGTGGTATTTACGGAAG 437  
 Oy 50 -----ThrIleGlnMetGly-----SerPhe 56  
 Db 438 AACCTCTATCAGACTTCTGGACTGTATGCTGTGGAAACAGAGGCCCACTGCTACTATTT 497  
 Oy 57 -----ArgIleAsnProAspGlySerGlnSer 65  
 Db 498 CCAGAGCCAGAAAGCAGCTCAGACCATGACAAAGAGGAAAGTATGATGGTTTGAGCACA 557  
 Oy 66 ValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGlu--- 84  
 Db 558 GAGCAGACGAGCGACCCCTTCATTTCCAGTGAAGAGTGCCTGTACACCTGTGTCCAAGGAAGT 617  
 Oy 85 ValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsn 104  
 Db 618 GCGCTGTATGAATTTGTTCTCTTAC-----CTCATTCAGAAAGTGAAGCAAGCAAAAT 671  
 Oy 105 TyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArg 124  
 Db 672 GTACTAGCGGCTGCTGTAGAACCTGAACTGATTTTGCATATGCCCATGCGAG----- 722  
 Oy 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGlnGluTyr 144  
 Db 723 -----GATATCAAGATGATCTCTAAATATGCTGACGCTGACCTCTATTTGAAGATTG 773  
 Oy 145 Glu 145  
 Db 774 GAA 776  
 RESULT 8  
 US-09-322-360-1  
 : Sequence 1, Application US/09322360  
 : Patent No. 6297050  
 : GENERAL INFORMATION:  
 : APPLICANT: Coulle, Pierre; Ikeda, Hideyuki;  
 : TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
 : TITLE OF INVENTION: Coding For Tumor Rejection Antigen Precursors DAGE and  
 : NUMBER OF SEQUENCES: 18  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Fulbright & Jaworski, L.L.P.  
 : STREET: 666 Fifth Avenue  
 : CITY: New York City  
 : STATE: New York  
 : COUNTRY: USA  
 : ZIP: 10103  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Diskette, 3.5 Inch, 1.44 MB storage  
 : COMPUTER: IBM PS/2  
 : OPERATING SYSTEM: PC-DOS  
 : SOFTWARE: Wordperfect  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/322,360  
 : FILING DATE:  
 : CLASSIFICATION:  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: 08/809,999  
 : FILING DATE:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Hanson, No. 6297050man D.

```

; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: IUD 5386.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: nucleic acid
; FEATURE:
; NAME/KEY: DAGE (5E10)
US-09-322-360-1

Alignment Scores:
Pred. No.: 0.638 Length: 1554
Score: 84.00 Matches: 40
Percent Similarity: 38.67% Conservative: 30
Best Local Similarity: 22.10% Mismatches: 61
Query Match: 8.76% Indels: 50
DB: 4 Gaps: 8

US-10-082-502-19 (1-182) x US-09-322-360-1 (1-1554)
OY 1 MetLysGlyTyrGlyTyrPLeuAlaLeuLeuGluValLeuLeuGlyThrAlaTPrLa 20
Db 276 GTGCGAGCCCTGCGCCCTTCACCTGCTCCCTCTGCGAGCTGTGAG----- 320
OY 21 ArgAArgSerGlnAAsPLeuHisCysGlyValAcuArgValLeuValAAsPLeuGlu--- 39
Db 321 ---AAGGACAAACATCTTCACCTGTGAGACCTTCAAAGCTGTGCTGAGTGAAGCATTGATGTC 377
OY 40 -----TPrLuiLeuAlaArgValAAsPProLysLys 49
Db 378 CTCCTTGCCCAAGAGGTTGCGCCCAAGAGGTGGAACCTTCAAGTGTGAGATTACGGAAG 437
OY 50 -----ThrIleGlnMetLys-----SerPhe 56
Db 438 AACTCTCATCAGACACTCTTGAGACTGTATGCTGTGGAACAGGCCACCTGTACTCATTT 497
OY 57 -----ArgIleAsnProAAsPArgLysSerGlnSer 65
Db 498 CCAGAGCCAGAACGACTCAGCCCATGACAAAGAACGAAAGTAGATGTTTGCAGCACA 557
OY 66 ValValGluValProTyrAlaIArgSerGlnLuiLHisLeuThrGluLeuGluGlu--- 84
Db 558 GAGGAGAGGACAGCCCTTCATTCACAGTAGAGGTGCTGTAGACCTGTCTTCAGGAAGGT 617
OY 85 ValCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThrHisArgLysAsn 104
Db 618 GCGCTGTGATGAATGTTCTCTCTAC-----CTCATTTGAAAGGAGCAGCAAGAAAAAT 671
OY 105 TyrValArgValIValSerArgAsnGlyGluSerSerGluLeuAsnPLeuGlnGlyIleArg 124
Db 672 GTACTACGCGGTGCTGTGAAGACTGAAGATTTTTCATGACGCCATGACAG----- 722
OY 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyr 144
Db 723 -----GATATTCAGATGATCTCTGAAAATGTGTGCAGCTGTATTTGAAGATTGG 773
OY 145 Glu 145
Db 774 GAA 776

RESULT 9
US-09-131-831B-1
; Sequence 1, Application US/09131831B
; Patent No. 6339149
; GENERAL INFORMATION:
; APPLICANT: Couille, Pierre; Ikeda, Hideyuki; Boon-
; Fallour, Thierry

```

TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
Coding For Tumor Rejection Antigen Precursors DAGE and  
Uses Thereof  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/131,831B  
FILING DATE: 11-Aug-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/809,999  
FILING DATE: 9-April-1997  
APPLICATION NUMBER: 08/316,231  
FILING DATE: 30-September-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6339149man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3100  
TELEFAX: (212) 318-3400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1554 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE (5E10)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-131-831B-1  
Alignment Scores:  
Pred. No.: 0.638 Length: 1554  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61  
Query Match: 8.76% Indels: 50  
Gaps: 8  
US-10-082-502-19 (1-182) x US-09-131-831B-1 (1-1554)  
QY 1 MetLygGlyTTPolYlTTPleuAlaLeuLeuLeuGlyYValLeuGlyThraAlaTTPala 20  
DB 276 GTGACAGCGCTGGCCCTGACCTGCTCCCTGCGAGTGTGATG----- 320  
QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlu--- 39  
DB 321 ---AAGGACACATCTTCCACCTGGAGACCTTAAAGCTGTGCTTGATGGACTTGATGTG 377  
QY 40 -----TTPGluIleAlaArgValAspProLyLyAs 49  
DB 378 CTCCTTGCCAGAGAGTGTGCCCCAGAGAGTGAACCTCAAGTGTGATTTACGGAAG 437  
QY 50 -----ThrlleGlnMetGly-----SerPhe 56  
DB 438 AACCTCATCAGAGACTTGTGGACTGTATGCTGTGAAACAGGCGCACTGTACTACTATT 497  
QY 57 -----ArgTleAsnProAspGlySerGlnSer 65  
DB 498 CCAGAGCCAGAGACGCTCAGCCCATGACAAAGAAAGGAAAGTAAGTGTGTTGAGACACA 557

QY 66 ValValGluValProTyrAlaIAspSerGluAlaHisLeuThrGluLeuLeuGluGlu--- 84  
DB 558 GAGGACAGACAGCCCTTCATTCACAGTAGAGTGTGCTGACSTGTCTCAAGGAAGT 617  
QY 85 ValCysAspArgMetCysGluTyrGlyGluGlnIleAspProSerThrHisArgLyAsn 104  
DB 618 GCCTGTGATGAATGTGTCTCTAC-----CTCATTAAGAAAGTGAAGCGAAAGAAAT 671  
QY 105 TyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyLeuArg 124  
DB 672 GTACTACGCTGTGCTGTGAAGAGCTGAAGATTTTGCAATGCCATGCCAG----- 722  
QY 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyr 144  
DB 723 -----GATATCAAGATGATCTGAAAGATGCTGACGTGACTGTATTGAAGATTG 773  
QY 145 Glu 145  
DB 774 GAA 776  
RESULT 10  
US-08-316-231B-1  
Sequence 1, Application US/08316231B  
Patent No. 5830753  
GENERAL INFORMATION:  
APPLICANT: Coullie, Pierre; Ikeda, Hideyuki;  
APPLICANT: Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules  
TITLE OF INVENTION: Coding For Tumor Rejection Antigen Precursors DAGE and  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Felte & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,231B  
FILING DATE: 30-September-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5830753man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5386  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
FEATURE:  
NAME/KEY: DAGE  
OTHER INFORMATION: Xaa is Arg when V is C or A or  
OTHER INFORMATION: Gly when V is G  
US-08-316-231B-1  
Alignment Scores:  
Pred. No.: 0.639 Length: 1555  
Score: 84.00 Matches: 40  
Percent Similarity: 38.67% Conservative: 30  
Best Local Similarity: 22.10% Mismatches: 61







Alignment Scores:			
Pred. NO.:	8.9e+04	Length:	4403765
Score:	82.50	Matches:	44
Percent Similarity:	40.41%	Conservative:	34
Best Local Similarity:	22.80%	Mismatches:	64
Query Match:	8.60%	Indels:	51
DB:	4	Gaps:	12

US-10-082-502-19 (1-182) x US-09-103-840A-2 (1-4403765)

OY 2 LysGIUTrGrIUTrLyeuAlaLeuLeuGluValLeuLeuGluTThrLaTrLaArg 21  
|||||  
Db 4382280 AAGGGCTGGGGGTGGGTGGCCSSSTATTATCGGAATGC-----TGCCGAC 4382233  
OY 22 --ArSerGIaAsrLeuNIscySGlValAcysuArgAlaLeuAlaSprLIuLeuGluTr 40  
|||  
Db 4382232 GGGGATTCGAAAGAGTAAGTGGGGGGC-----GTGCCGTGG 4382194  
OY 41 -----GlUIleAlaArgValAsprOLyLvsThIle---GlmeIGly 54  
:::  
Db 4382193 CGSAGCCGGGGCAAGCTGCGSSTACAGACTTCAGTCSAGSCAGCGGGGGCGATGGCC 4382134  
OY 55 SerPheArgIIeaInProARPrLyserLIeBerValAlaGluValProTyAlaArgSer 74  
|||  
Db 4382133 GTATTGGCCCCAGACAGCTCCGGAGCGGGGGGTGATCAAGTGGGGTTGTGACGGC 4382074  
OY 75 GlUAInHIsLeuThrGluLeuLeu-----GluGluValAcysAPrgMetLysGIu 91  
|||  
Db 4382073 TAGCCAGACACACAGCTGTCTGTGGCCAATGCCACACAGCTGGCGCCAGCTAGACGC 4382014  
OY 92 TyrcLgLuGlnLIleaRPrOSerThrlaArgLysAbnTyValaArgValaLseArg 111  
|||  
Db 4382013 TATGCCCGCCGATTCGAC--CGGGTGCAT-----4381987  
OY 112 AsnGLyUsrSeSGluLeuAspLeuIncLYIleArgLIleAsrSerApLIeserGly 131  
:::::  
Db 4381986 -----GCGGGGTCTCGGATTTGTGGC--CGCATCTCGCATCCGCTGACCGGG 4381939  
OY 132 ThrLeuSPheaLIcysGILserLIeValGluGluTyGluAspGILue-----148  
::::|  
Db 4381938 ---ATCAAAGAAGTGTGGAGTTTCTACAGCACAGACAGAACAGATCSAGGCATC 4381882  
OY 149 -----IleGUlBheSerArgGluAlaAspValaLysAr 161  
:::::  
Db 4381881 GCCCATGACATCGCGGTGTGTGACACATTCAGCGCGGAGAGN-----GAC 4381834  
OY 162 LysLeucysSerLyArgTrhAspLeucysAspHISAla 174  
|||  
Db 4381833 GCCTTGGCTCGGAGATCACCCGCGTGTGTGCACGCC 4381795  
  
RESULT 15  
US-09-103-840A-1/c  
Sequence 1, Application US/09103840A  
Patent No. 6294328  
GENERAL INFORMATION:  
APPLICANT: FLEISCHMAN, Robert D.  
APPLICANT: WHITE, Owen R.  
APPLICANT: FRASER, Claire M.  
APPLICANT: VENTER, John C.  
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
FILE REFERENCE: 24366-2007.00  
CURRENT APPLICATION NUMBER: US/09/103,840A  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 4411529  
TYPE: DNA  
ORGANISM: Mycobacterium tuberculosis  
OTHER INFORMATION: H37rv  
US-09-103-840A-1

## Alignment Scores:

Pred. No.:	8.93e+04	Length:	4411529
Score:	82.50	Matches:	44
Percent Similarity:	40.41%	Conservative:	34
Best Local Similarity:	22.80%	Mismatches:	64
Query Match:	8.60%	Indels:	51
DB:	4	Gaps:	12

US-10-082-502-19 (1-182) x US-09-103-840A-1 (1-4411529)

```
OY      2  LysGlyTrpGlyTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAlaArg 21
          |||||
Db 4390042 AAGGGCTGGGGTGGGGTGGCCGATATCGGGATGATC-----TGCCCGAAC 4389995
          |||||

OY      22  ---ArgSerGluAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40
          |||
Db 4389994 GGGGATTGCAAAAGCTACGTCCGGGGCT-----CTGGCGTGG 4389956
          |||

OY      41  -----GluIleAlaArgValAspProLysLysThrIle---GlnMetGly 54
          ::|||
Db 4389955 CGCACCGCGGGCCAGTTCGCGCTGACTGACTGATTCAGTTCAGCGCGCGCGGCGGATGGCG 4389896
          ::|||

OY      55  SerPheArgIleAsnProAspArgLysSerGlnSerValValGluValProTyraIleArgSer 74
          |||
Db 4389895 GTTATTCCGCCGCCACACGCTCCCGAGCGGGCGGTGATCGAGTCGCGCGCTTGTCTGACGCG 4389836
          |||

OY      75  GluAlaHisLeuThrGluLeuLeu-----GluGluValCysAspArgMetLysGlu 91
          |||
Db 4389835 TACGCCACACACCGCGCTGTCTGGGCCAATCCACACGCTGCGGCCCGCCAGCTAGACGCC 4389776
          |||

OY      92  TyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyValAlaArgValAlaSerArg 111
          |||
Db 4389775 TATGCCGCCGATCGAC---CGGGTGCAT----- 4389749
          |||

OY      112 AsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAspIleSerGly 131
          ::|||
Db 4389748 -----GCCGCGCTGCTGATTTGTTGGCC---CGCATCTGCGATCCGCTGACCGGG 4389701
          ::|||

OY      132 ThrLeuLysPheAlaCysArgLysSerIleValGluGlyIleValAspGluLeu----- 148
          ::|||
Db 4389700 ---ATCAAGAGAGCTGTGGAGCTTCTGACGCCACAGACAGAGACAGATCCAGCCGATC 4389644
          ::|||

OY      149  -----IleGluPhePheSerArgGluAlaAspAsnValLysAsp 161
          ::|||
Db 4389643 GCCCATGACATCGCGGTGGTGTGACCACTTCAGCGGGAAGTG-----GAC 4389596
          ::|||

OY      162 LysLeuCysSerLysArgThrAspLeuCysAspHisAla 174
          |||
Db 4389595 CGGTGGCTGGCGAGATCACCGGGTGGTGTGTCGACGCC 4389557
          |||
```

Search completed: March 31, 2003, 04:26:04  
Job time : 1346.68 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:08:03 ; Search time 142.849 Seconds

(without alignments)  
2869.213 Million cell updates/sec

Title: US-10-082-502-19  
Perfect score: 959

sequence: 1 MNGMGWLALLLVLLGTAWA.....LCSKRTDLCDHALHRSDEL 182

**Scoring table:**

Ygapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
delop	6.0	delext	7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Maximum DB seq length: 20000000000

Post-Processing: Minimum Match 0%

### Listing first 45 summaries

Command line parameters:  
-MODEL=fract 333 444 555

[illegible]

Database :

1:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1960.DAT*
2:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1961.DAT*
3:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1962.DAT*
4:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1963.DAT*
5:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1964.DAT*
6:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1965.DAT*
7:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1966.DAT*
8:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1967.DAT*
9:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1968.DAT*
10:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1969.DAT*
11:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1970.DAT*
12:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1971.DAT*
13:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1972.DAT*
14:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1973.DAT*
15:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1974.DAT*
16:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1975.DAT*
17:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1976.DAT*
18:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1977.DAT*
19:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1978.DAT*
20:	/SDS2/gcgdata/genseq/genseqn-emb1/NA1979.DAT*
21:	/SDS2/gcgdata/genseq/genseqn-emb1/NA2000.DAT*
22:	/SDS2/gcgdata/genseq/genseqn-emb1/NA2001.DAT*
23:	/SDS2/gcgdata/genseq/genseqn-emb1/NA2001A.DAT*
24:	/SDS2/gcgdata/genseq/genseqn-emb1/NA2001B.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES						Description
Result No.	Score	% Query Match	Length	DB	ID	
1	959	100.0	1069	20	AAK06970	Mouse secretory p
2	959	100.0	1069	21	AAZ08294	Mouse ortholog p
3	938	97.8	546	21	AAZ38326	Human transmembran
4	938	97.8	806	20	AAK06969	Human secretory p
5	938	97.8	806	21	AAZ08293	Human Zs19 gene
6	938	97.8	814	21	AAZ38327	Human transmembran
7	938	97.8	814	24	ABK52765	cDNA encoding tra
8	938	97.8	814	24	ABL41995	Nucleotide sequen
9	938	97.8	814	24	ABK09772	Human ovarian tum
10	938	97.8	1180	20	AAZ22112	Human secreted pri
11	938	97.8	1210	22	AAZ21330	Human cDNA sequen
12	931	97.1	714	20	AAZ97884	Human secreted pri
13	895	93.4	657	24	ABK36007	cDNA sequence #39
14	889	92.7	832	20	AAZ97837	Human secreted pri
15	856.5	89.3	592	21	AAA43525	Mouse secreted ex
16	380	39.6	484	22	AAL25628	Human breast canc
17	380	39.6	558	22	AAL18023	Human breast canc
18	380	39.6	595	22	AAL07882	Human breast canc
19	380	39.6	1160	22	AAL26685	Human secretory p
20	361	37.6	649	20	AAK57863	Human secreted pri
21	361	37.6	649	20	AAK06968	Human secretory p
22	361	37.6	649	21	AAZ08284	Human secreted pri
23	357	37.2	718	20	AAZ10655	CDNA encoding a h
24	357	37.2	792	20	AAZ97836	Human secreted pri
25	357	37.2	415	21	AAZ08285	Human secreted pri
26	347	36.2	1085	22	AAZ2130	Expressed Sequence
27	347	36.2	1085	22	AAZ2130	WO9901020 Seq ID
28	308	32.1	2532	22	ABA07341	Human pancreatic t
29	308	32.1	2724	22	ABA07340	Human pancreatic t
30	308	32.1	2724	22	ABA07340	Human genomic DNA
31	308	32.1	2724	22	ABA07340	Human genomic DNA
32	262	27.3	564	24	ABK09691	Human ovarian tum
33	212.5	22.2	816	23	ABL21471	Drosophila melano
34	180	18.8	352	22	ABA09488	Human secreted pri
35	157	16.4	785	21	AAZ37767	Arabidopsis thalli
36	145	15.1	750	23	ABL16485	Drosophila melano
37	145	15.1	750	23	ABL16484	Drosophila melano
38	145	15.1	750	23	ABL16484	Drosophila melano
39	140	14.6	3110	23	ABL21470	Drosophila melano
40	136.5	14.2	744	21	AAZ96104	Human myocardi
41	136.5	14.2	744	21	AAZ96104	Ptarm specific fct
42	136.5	14.2	747	22	AAZ90704	Human secretory pr
43	136.5	14.2	1419	22	AAZ93314	Human cDNA encod
44	136.5	14.2	1457	20	AAZ78923	Human calca-1 DNA
45	136.5	14.2	1462	21	AAA16621	Human secreted pri

ALIGNMENTS

RESULT 1  
ID AAK06970 standard; CDNA; 1069 BP.  
AC AAX06970;  
DT 10-MAY-1999 (first entry)  
DE Mouse secretory peptide-9 (Zs19) orthologue CDNA.  
KW Secretory peptide-9; Zs19; orthologue; mouse; tumour marker;  
XX cancer; therapy; diagnosis; growth enhancer; ss.  
OS Mus sp.  
XX  
XX  
XX Key Location/Qualifiers  
FT CDS /tag= a  
FT 358..906  
FT

```

FT      sig_peptide      358..417
FT      /*tag= b
FT      mat_peptide      418..903
FT      /*tag= c
XX
XX      MO9901554-A1.
XX
XX      14-JAN-1999.
XX
XX      02-JUL-1998;      98WO-US13859.
XX
XX      17-JUN-1998;      98US-0099005.
XX      03-JUL-1997;      97US-0051704.
XX      03-JUL-1997;      97US-0888088.
XX      19-MAY-1998;      98US-0081338.
XX      19-MAY-1998;      98US-0085983.
XX      17-JUN-1998;      98US-0089899.
XX
XX      (ZYMO ) ZYMOGENETICS INC.
XX
XX      Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;
XX
XX      WPI: 1999-106055/09.
XX      P-PSDB: AAM88476.
XX
XX      New mammalian secretory peptide-9 (zsig9) - used as a growth
XX      enhancer for placenta, liver and heart, and as an indicator of
XX      cancer
XX
XX      Claim 2; Page 75-77; 85pp; English.
XX
XX      This CDNA clone encodes novel mouse secretory peptide-9, or zsig9
XX      (see AAM88476), an orthologue of novel human zsig9 (see AAM88469).
XX      Human zsig9 is overexpressed in a number of tumours including
XX      brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid
XX      and lymphoma tumors, and thus can be used as an indicator for
XX      cancer. The invention provides polynucleotides (see AAM06968-70)
XX      encoding zsig9 polypeptides (see AAM88469-77) including mature
XX      polypeptides, other processed forms, variants and the mouse
XX      orthologues. The zsig9 gene, or probes derived from it, can be
XX      used to determine if zsig9 is present on chromosome 10, and if a
XX      mutation has occurred. Antibodies raised against zsig9 can be
XX      used as diagnostic agents to determine the presence of zsig9, and
XX      thus the presence of cancer. They can also be labelled with
XX      radioisotopes or fused with toxins and used to treat tumours
XX      which overexpress zsig9. Antisense nucleotides derived from zsig9
XX      cDNA can also be used to inhibit the growth of tumour cells. zsig9
XX      proteins can be used to enhance the growth or development of the
XX      placenta, heart or liver.
XX
XX      Sequence 1069 BP: 276 A; 259 C; 325 G; 209 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. NO.:      8.05e-102      Length:      1069
XX      Score:      959.00      Matches:      182
XX      Percent Similarity:      100.00%      Conservative:      0
XX      Best Local Similarity:      100.00%      Mismatches:      0
XX      Query Match:      100.00%      Indels:      0
XX      DB:      20      Gaps:      0
XX
XX      US-10-082-502-19 (1-182) x AAM06970 (1-1069)
XX
XX      QY      1 MetLysGlyTyrGlyTyrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrrPa 20
XX      DB      358 ATGAAGAGCTGGGGTGGCTAGCCCTACTTGTGGGGTCTGCTGGAACTGGGCT 417
XX
XX      QY      21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGlnTrr 40
XX      DB      418 CGAAGGAGCCCAAGATCACTGAGCTGGAGCTTGCAAGGCTCGGAGATTAAGAGTG 477
XX
XX      QY      41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60
XX      DB      478 GAAATTCGCCCGCTGGAGCCCAAGAGACATTCAGATGGGATCTTCGATCAATCA 537

```

```

QY      61 AspGlySerGlnSerValValGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80
DB      538 GATGGCAGCAGCTCAGTTGTGGAGGTACTTATGCCCGCTGAGAGCCACCTCAGAG 597
QY      81 LeuLeuGluGluValCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr 100
DB      598 TTGCTTGAGAGGTGTGTGAGCGAATGAGAGTGGGGAACAGATTGACCTTCTACG 657
QY      101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 120
DB      658 CACCCCAAGAACTAGCTAGCTACGCTCTCTGAGCCGGAATGGAGAACTCAGAACTAGACTTA 717
QY      121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
DB      718 CAGGCAATCCGAAATGACTCAGATATCAGCGGACCCCTCAAGTTGCGTGTAGAGCAT 777
QY      141 ValGluGluTyrGluAspGluLeuLeuIleGluPheSerArgGluAlaAspAsnValLys 160
DB      778 GTGGAAGATACGAGAGATGAGCTTATCGAATTCTTCCAGAGAGGCTGACAAACCTTAAA 837
QY      161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
DB      838 GACAAACTTTGCAGTAACCGGACAGACATCTATGTGACCATGCCCTGCACAGATCTCAGCAT 897
QY      181 GluLeu 182
DB      898 GAGCTA 903

RESULT 2
ID      AAZ08294
XX      AAZ08294 standard; DNA; 1069 BP.
XX
XX      AAZ08294;
XX
XX      07-FEB-2000 (first entry)
XX
XX      Mouse ortholog gene encoding zsig9 secretory protein.
XX
XX      Secretory protein-9; mouse zsig9; ortholog; overexpression; antagonist;
XX      antibody; antisense nucleotide; tumour; treatment; receptor;
XX      radio-label; polypeptide toxin; down-regulation; diagnostic;
XX      therapeutic; probe; cancer; brain; liver; detection; stomach;
XX      lymphoma; ds.
XX
XX      Mus musculus.
XX
XX      Key      Location/Qualifiers
XX      CDS      358..906
XX
XX      FT      /*tag= a
XX      FT      /product= "Mouse zsig9 secretory protein"
XX      FT      /note= "Overexpressed in tumours"
XX      FT      sig_peptide      358..417
XX      FT      mat_peptide      418..903
XX      FT      /*tag= b
XX      FT      /*tag= c
XX      FT      /label= Mature_zsig9_protein_variant-4
XX
XX      MO9960405-A1.
XX
XX      PD      25-NOV-1999.
XX
XX      XX      19-MAY-1999;      99WO-US11107.
XX      PF      19-MAY-1999;      98US-0081183.
XX      PR      19-MAY-1998;      98US-0081183.
XX
XX      XX      (ZYMO ) ZYMOGENETICS INC.
XX      PA      Moore EE, Taft DW;
XX      PI      WPI: 2000-039447/03.
XX      DR      P-PSDB: AAI15136.
XX
XX

```

PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (Zs19)  
 XX  
 XX  
 PS Disclosure; Page 37-38; 45pp; English.  
 CC The present DNA sequence is an ortholog encoding the secretory protein-9,  
 CC Zs19 derived from mouse. It is overexpressed in tumours. Antagonists,  
 CC antibodies and antisense nucleotides to Zs19 are useful for detecting  
 CC and treating tumours. The antagonist may be an antibody or receptor to  
 CC Zs19 and it may be radio-labelled or fused to a polypeptide toxin. It  
 CC can be used for down regulating the overexpression of Zs19. The gene  
 CC sequence can be used as nucleic acid probes to detect RNA encoding Zs19.  
 CC The Zs19 sequence facilitates improved diagnostic and therapeutic  
 CC techniques for detecting and treating cancers, especially of the brain,  
 CC liver, stomach, lymphoma etc., at an early stage.  
 CC  
 XX  
 SO Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 8,05e-102 Length: 1069  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0  
 US-10-082-502-19 (1-182) x AA208294 (1-1069)  
 OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 358 ATGAAAGGCTGGGCTGGCTGACCCCTACTTTTGGGGCTCCTGGGAACTGGCTGGCT 417  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 418 CGAAGAGAGCCAAAGATCTACACTGTGGAGCTTGAGAGGCTGTGGATGAATTAGATGG 477  
 OY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 478 GAAATGCGCCGCTGGACCCCAAGAACCATTCAGATGGGATCTCCGAATCAATCA 537  
 OY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisIleuThrGlu 80  
 DB 538 GATGGAGCCAGTCAGTGTGGAGGTACCTTATGCCCGCTCAAGGCCCACTCAGAG 597  
 OY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGluIleAspProSerThr 100  
 DB 598 TTCCTGAGGAGGCTGTGTGACCGAATGAAGAGTACGGGAACAGATTGACCTTCTAC 657  
 OY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 658 CACCCGAAAGAACTACGACCGCTGAGCCGGAATGGAATCACTGAACTAGACTTA 717  
 OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 718 CAGGGATCCGAATGTGACTAGATATCAGCGGACCCCTCAAGTTTGGGTGAGAGCAT 777  
 OY 141 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB 778 GTGGAAAGAATACAGAGATGAGCTTATCGATTCTTCCAGAGAGCGCTGCAACGTTAA 837  
 OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 838 GACAAACTTTGACGTAAGCGGACAGATCTATGTGACCATGCCCTGCACAGATCTCAGAT 897  
 OY 181 GluLeu 182  
 DB 898 GAGCTA 903  
 RESULT 3  
 AA238326  
 ID AA238326 standard; cDNA; 546 BP.  
 XX  
 AC AA238326;

XX  
 XX 09-FEB-2000 (first entry)  
 XX  
 XX Human transmembrane protein CDNA clone HP10390 coding sequence.  
 DE  
 DE  
 KW HP10390; transmembrane domain; stomach cancer cell; antibody;  
 KW assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;  
 KW agonist; antagonist; ligand; therapeutic; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH CDS 1..540  
 FT /tag- a  
 FT /product- "Human transmembrane protein HP10390"  
 FT /note- "No stop codon given in the specification"  
 FT  
 PN MO9955862-A2.  
 XX  
 PD 04-NOV-1999.  
 XX  
 PF 27-APR-1999; 99MO-JP02226.  
 XX  
 PR 28-APR-1998; 98JP-0119395.  
 XX  
 PA (SAGA ) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;  
 XX  
 DR WPI; 2000-023358/02.  
 DR P-PSDB; AAY52391.  
 XX  
 PT Human proteins with transmembrane domains, involved in control of cell  
 PT proliferation and differentiation, useful for treating e.g. cancer or  
 PT inflammation  
 XX  
 PS claim 3; Page 88; 114pp; English.  
 CC  
 CC This sequence represents the coding sequence of human CDNA clone  
 CC HP10390 which encodes a 20 kD protein with one putative transmembrane  
 CC domain in the N-terminus. The CDNA was isolated from a human stomach  
 CC cancer cell line CDNA library. The protein has no homology with any  
 CC known protein. The protein may be used to raise specific antibodies, as  
 CC assay reagents, as diagnostic tissue markers, for the isolation of  
 CC cognate receptors, ligands and binding proteins, and as biologically  
 CC active agents. Nucleotides encoding the protein may be used as primers  
 CC and probes or antisense molecules, and in gene therapy. Cells transformed  
 CC with these nucleotides may be used to screen for agonists and antagonists  
 CC which are potentially useful therapeutically.  
 CC  
 SO Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 9,09e-100 Length: 546  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 21 Gaps: 0  
 US-10-082-502-19 (1-182) x AA238326 (1-546)  
 OY 1 MetLysGlyTrpGlyTrpLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 1 ATGAAAGGCTGGGCTGGCTGACCCCTACTTTTGGGGCTCCTGGGAACTGGCTGGCT 60  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 61 CGAGAGAGCCAGATCTCCACTGTGGAGCATGCGAGGCTGTGGATGAATGAATGG 120  
 OY 41 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60



```

Db 121 GAATTGCCAGTGGAGCCCAAGAACATTCAGATGGAGTCTTTCCGATCAATCCA 180
Oy 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80
Db 181 GATGCGACGACGACTCAGTGTGGAGGTGCTTATGCCCGCTGAGAGCCGACCTCAGAG 240
Oy 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnHisLeuProSerThr 100
Db 241 CTGCTGGAGGAGATATGTCACCGGATGGAAGGATGGGGAACATATGATCTTCCACC 300
Oy 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 120
Db 301 CATCCCAAGAACTACTACGTACGTACTGGGCGGAAATGGAGATCCAGTCACTGAACTGA 360
Oy 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140
Db 361 CAAGGCATCCGAAATGAGTCAAGATATATAGCGGCACCCCTCAAGTTTGGCTGTAGAGCAT 420
Oy 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160
Db 421 GTGGAGGAATACGAGATGAGAACTCATTAATCTTTCCCGAGGCTGACAAATGTTAA 480
Oy 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180
Db 481 GACAAACCTTTCAGTAAAGCAACACATCTTTGTGACCATGCCCTGCACATATCGCATGAT 540
Oy 181 GluLeu 182
Db 541 GAGCTA 546

```

```

PT New mammalian secretory peptide-9 (Zs19) - used as a growth
PT enhancer for placenta, liver and heart, and as an indicator of
PT cancer
PS Claim 2; Page 73-74; 85pp; English.
XX
CC This cDNA clone encodes human secretory peptide-9, or Zs19,
CC variant (see AAM88474). Zs19 (see also AAM88469) is overexpressed in
CC human brain, liver, lung, oesophageal, stomach, colon, rectal,
CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an
CC indicator for cancer. Zs19 cDNA was discovered in a placenta
CC clone from a full-term pregnancy cDNA library which contained an
CC expressed sequence tag (see AAX06971). The invention provides
CC polynucleotides (see AAX06968-70) encoding Zs19 polypeptides (see
CC AAM88469-77) including mature polypeptides, other processed forms,
CC variants and mouse orthologues. The Zs19 gene, or probes derived
CC from it, can be used to determine if Zs19 is present on chromosome
CC 12, and if a mutation has occurred. Antibodies raised against
CC Zs19 can be used as diagnostic agents to determine the presence of
CC Zs19, and thus the presence of cancer. They can also be labelled
CC with radioisotopes or fused with toxins and used to treat tumors
CC which overexpress Zs19. Antisense nucleotides derived from Zs19
CC cDNA can also be used to inhibit the growth of tumour cells. Zs19
CC proteins can be used to enhance the growth or development of the
CC placenta, heart or liver.
XX
SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other:

```

```

Alignment Scores:
Pred. No.: 1,53e-99 Length: 806
Score: 938.00 Matches: 177
Percent Similarity: 98.35% Conservative: 2
Best Local Similarity: 97.25% Mismatches: 3
Query Match: 97.81% Indels: 0
DB: Gaps: 0

```

US-10-082-502-19 (1-182) x AAX06969 (1-806)

```

DE 10-MAY-1999 (first entry)
XX
DE 10-MAY-1999 (first entry)
XX
DE Human secretory peptide-9 (Zs19) variant cDNA.
XX
KW Secretory peptide-9; Zs19; human; tumour marker; cancer; therapy;
XX diagnosis; growth enhancer; variant; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS /*tag= a
XX FT 104..652
XX FT sig_peptide /*tag= b
XX FT 104..163
XX FT mat_peptide /*tag= c
XX FT 164..649
XX FT /*tag= c
XX
PN W09901554-A1.
XX
XX 14-JAN-1999.
XX
XX 02-JUL-1998; 98WO-US13859.
XX
XX 17-JUN-1998; 98US-0099005.
XX 03-JUL-1997; 97US-0051704.
XX 03-JUL-1997; 97US-0888088.
XX 19-MAY-1998; 98US-0081338.
XX 19-MAY-1998; 98US-0085983.
XX 17-JUN-1998; 98US-0089899.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;
XX WPI: 1999-106055/09.
XX P-PSDB; AAM88474.
XX

```

```

PT New mammalian secretory peptide-9 (Zs19) - used as a growth
PT enhancer for placenta, liver and heart, and as an indicator of
PT cancer
PS Claim 2; Page 73-74; 85pp; English.
XX
CC This cDNA clone encodes human secretory peptide-9, or Zs19,
CC variant (see AAM88474). Zs19 (see also AAM88469) is overexpressed in
CC human brain, liver, lung, oesophageal, stomach, colon, rectal,
CC thyroid and lymphoma tumors. Thus, Zs19 can be used as an
CC indicator for cancer. Zs19 cDNA was discovered in a placenta
CC clone from a full-term pregnancy cDNA library which contained an
CC expressed sequence tag (see AAX06971). The invention provides
CC polynucleotides (see AAX06968-70) encoding Zs19 polypeptides (see
CC AAM88469-77) including mature polypeptides, other processed forms,
CC variants and mouse orthologues. The Zs19 gene, or probes derived
CC from it, can be used to determine if Zs19 is present on chromosome
CC 12, and if a mutation has occurred. Antibodies raised against
CC Zs19 can be used as diagnostic agents to determine the presence of
CC Zs19, and thus the presence of cancer. They can also be labelled
CC with radioisotopes or fused with toxins and used to treat tumors
CC which overexpress Zs19. Antisense nucleotides derived from Zs19
CC cDNA can also be used to inhibit the growth of tumour cells. Zs19
CC proteins can be used to enhance the growth or development of the
CC placenta, heart or liver.
XX
SQ Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other:

```

QY 181 GIULEU 182  
 Db 644 GAGCTA 649  
 RESULT 5  
 AAZ08293  
 ID AAZ08293 standard; DNA; 806 BP.  
 AC AAZ08293;  
 XX  
 DT 07-FEB-2000 (first entry)  
 XX  
 DE Human zsi9 gene encoding secretory protein variant-4.  
 XX  
 KW Secretory protein-9; Human zsi9; chromosome 12q15 region; variant;  
 KW overexpression; antagonist; antibody; antisense nucleotide; tumour;  
 KW treatment; receptor; radio-label; fusion; polypeptide toxin; technique;  
 KW down-regulation; probe; diagnostic; therapeutic; cancer; brain; liver;  
 KW detection; stomach; lymphoma; alternative splicing; allelic variation;  
 KW silent mutation; ds.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 104..652  
 FT /tag- a  
 FT /product- "zsi9 secretory protein variant-4"  
 FT /note- "Overexpressed in tumours"  
 FT 104..163  
 FT /tag- b  
 FT 164..649  
 FT /tag- c  
 FT /label- Mature\_zsi9\_protein\_variant-4  
 FT  
 XX  
 PN W09960405-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 19-MAY-1999; 99MO-US11107.  
 XX  
 PR 19-MAY-1998; 98US-0081183.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Moore EE, Taft DW;  
 XX  
 DR WPI: 2000-039447/03.  
 DR P-PSDB: AAY13135.  
 XX  
 PT Detecting tumors using antibodies, antagonists and antisense  
 PT nucleotides to secretory protein-9 (zsi9) -  
 XX  
 PS Disclosure: Page 33-35; 45pp; English.  
 XX  
 CC The present DNA sequence is a gene encoding the variant of the secretory  
 CC protein-9, zsi9 that arises due to alternative splicing, allelic  
 CC variation or silent mutations that result in amino acid changes. This  
 CC sequence is mapped to the human chromosome 12q15 region. It is  
 CC overexpressed in tumours. Antagonists, antibodies and antisense  
 CC nucleotides to zsi9 are useful for detecting and treating tumours. The  
 CC antagonist may be an antibody or receptor to zsi9 and it may be radio-  
 CC labeled or fused to a polypeptide toxin. It can be used for down  
 CC regulating the overexpression of zsi9. The gene sequence can be used as  
 CC nucleic acid probes to detect RNA encoding zsi9. The zsi9 sequence  
 CC facilitates improved diagnostic and therapeutic techniques for detecting  
 CC and treating cancers, especially of the brain, liver, stomach, lymphoma,  
 CC etc., at an early stage.  
 XX  
 SQ Sequence 806 BP; 229 A; 183 C; 233 G; 160 T; 1 other;  
 Alignment Scores:

Pred. No.: 1,53e-99 Length: 806  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 21 Gaps: 0  
 US-10-082-502-19 (1-182) x AAZ08293 (1-806)  
 QY 1 MetysgllytrpglytrpleuAlaLeuLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 Db 104 ATGAAGGCTGGGGTGGTGGCCCTCTCTGGGGGCCCTGGCGGAACCGCTGGCT 163  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaLcysArgAlaLeuValAspGlyLutp 40  
 Db 164 CGGAGGAGCCAGAGATCTCCAGTGGAGCAATGAGAGGCTCTGGTGAATGAAATGAG 223  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 Db 224 GAATATGCCAGGTGGACCCCAAGAGACCATTCAGATGGGATCTTCGGATCAATCCA 283  
 QY 61 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80  
 Db 284 GATGGCAGCCAGTCAGTGGTGGAGTGGCTTATGCCCGTCAGAGGCCACCTCACAGAG 343  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 100  
 Db 344 CTCTGGAGGAGATATGTGACCGGATGMAAGATATGAGAACATGATCTCTCCACC 403  
 QY 101 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 Db 404 CATCGCAGAACTACCTAGCTGTAAGTGGCCGGAATGGAATCCAGTCACTGAGCTTA 463  
 QY 121 GlnGlyIleArgThrAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 Db 464 CAAAGCATCCGAATCGACATCAGATATATACCGGACCCCTCAAGTTCCTGTCAGAGCAT 523  
 QY 141 ValGluGluTyrGluAspLeuLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 Db 524 GTGAGAGAAATACAGAGATGACATGATGATCTTTCCGAGAGCTACAAATGTAA 583  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 Db 584 GACAAACTTTCAGATGAAGCAACAGATCTTGTGACCATGCCCTGCACATATGCATGAT 643  
 QY 181 GIULEU 182  
 Db 644 GAGCTA 649  
 RESULT 6  
 AAZ38327  
 ID AAZ38327 standard; cDNA; 814 BP.  
 AC AAZ38327;  
 XX  
 DT 09-FEB-2000 (first entry)  
 XX  
 DE Human transmembrane protein cDNA clone HP10390.  
 XX  
 KW HP10390; transmembrane domain; stomach cancer cell; antibody;  
 KW assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;  
 KW agonist; antagonist; ligand; therapeutic; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 145..693  
 FT /tag- a  
 FT /product- "Human transmembrane protein HP10390"  
 FT  
 XX  
 PN W09955862-A2.  
 PD 04-NOV-1999.



CC animal, where the composition elicits the production of cytotoxic T  
 CC lymphocytes specific for the immunogenic composition. Preferably, the  
 CC animal is a human. The cancer-linked genes and polypeptides are also  
 CC useful as targets for cancer therapy or chemotherapy. The present  
 CC sequence represents a cancer-linked gene located on chromosome 12,  
 CC which encodes transmembrane protein 4.

Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:

Pred. No.: 1.55e-99 Length: 814  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 24 Gaps: 0

US-10-082-502-19 (1-182) x ABK52765 (1-814)

QY 1 MetlysglytrpGLYTPRLPLeuAlaLeuLeuGLYValLeuLeuGLYThrAlaTrpAla 20  
 DB 145 ATGAAGAGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 205 CGAGAGAGCCAGAGATCTCAGCTGGAGCATGAGAGGCTCTGGTGGATGAATGAAATGG 264  
 QY 41 GluLeuAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 265 GAAATGCGCCAGGTGGACCCCAAGAGACATTCAGATGGATCTTCCGATCAATCCA 324  
 QY 61 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 325 GATGGACAGCCAGTGGCTGGAGTGGCTTATGCTGGCTGGCTGGCTGGCTGGCTGGCT 384  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 385 CTGCTGGAGAGATATGTGACCGGATGAGATGAGATGAGATGAGATGAGATGAGATG 444  
 QY 101 HisArgLysAsnTrpValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 445 CATGGCAAGAACTACGATCGATGAGTGGCGGCAATGAGAACTGAACTGAGACTTA 504  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 505 CAAGGCATCCGAATCGACTGAGATATTAGCGGCACCTCAAGTTGGCTGGAGAGCAT 564  
 QY 141 ValGluGluTrpGlyValAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 DB 565 GTGGAGGAATACGAGATGAACTCATTTCTTCCCGAGAGGCTGACAAATGTTTAA 624  
 QY 161 AspLysLeuCysSerLysArgTrpAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 625 GACAAACTTTGCAAGTAAAGCAAGATCTTTGACCATCCCTCGACATATGTCAGAT 684  
 QY 181 GluLeu 182  
 DB 685 GAGCTA 690

RESULT 8  
 ID ABL41995 standard; DNM: 814 BP.

AC ABL41995;

DE 11-JUN-2002 (first entry)

XX Nucleotide sequence of human polypeptide HP10390.

XX Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers  
 CDS 145..693  
 /tag= a  
 /product= "HP10390"

MO200208416-A1.

31-JAN-2002.

24-JUL-2001; 2001WO-JP06371.

24-JUL-2000; 2000JP-0222743.

24-AUG-2000; 2000JP-0254407.

(NISC-) JAPAN SCI & TECHNOLOGY CORP.

Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;

WPI; 2002-195877/25.

P-PSDB; ABB09717.

Antibody preparation by inoculation of an animal with a vector  
 expressing a fusion protein of an antigen on the C-terminal side of a  
 transmembrane domain for use as drugs, diagnostic reagents and  
 laboratory reagents

Example; Page 33-35; 45pp; Japanese.

The specification describes a method of antibody preparation. The  
 method comprises inoculating an animal with a vector expressing  
 a fusion protein having an antigen protein fused to the C-terminal  
 side (extracellular) of a transmembrane domain protein (the  
 N-terminal side of which is intracellular), and then isolating and  
 purifying the antibody from the animal. The antibodies can be used  
 as drugs, vaccines, diagnostic reagents and laboratory reagents. The  
 present sequence encodes a polypeptide, designated HP10390, which was  
 used in the course of the invention.

Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:

Pred. No.: 1.55e-99 Length: 814  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: 24 Gaps: 0

US-10-082-502-19 (1-182) x ABL41995 (1-814)

QY 1 MetlysglytrpGLYTPRLPLeuAlaLeuLeuGLYValLeuLeuGLYThrAlaTrpAla 20  
 DB 145 ATGAAGAGCTGGGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 204  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB 205 CGAGAGAGCCAGAGATCTCAGCTGGAGCATGAGAGGCTCTGGTGGATGAATGAAATGG 264  
 QY 41 GluLeuAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 265 GAAATGCGCCAGGTGGACCCCAAGAGACATTCAGATGGATCTTCCGATCAATCCA 324  
 QY 61 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 325 GATGGACAGCCAGTGGCTGGAGTGGCTTATGCTGGCTGGCTGGCTGGCTGGCTGGCT 384  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 100  
 DB 385 CTGCTGGAGAGATATGTGACCGGATGAGATGAGATGAGATGAGATGAGATGAGATG 444  
 QY 101 HisArgLysAsnTrpValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 120  
 DB 445 CATGGCAAGAACTACGATCGATGAGTGGCGGCAATGAGAACTGAACTGAGACTTA 504

OY 121 GlnGlyIleAArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 505 CAAAGCATCCGAAATCGACTAGATATAGCGCACCCCTCAAGTTGCGGTGAGAGCAT 564  
OY 141 ValGluGluTyrGlnAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 555 GTGAGAGAAATACGAGATGAACCTCATTAATCTTTCCCGAGAGCTGACAAATGTTAA 624  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
DB 625 GACAAACTTTCAGATACGAAACAGATCTTTGTGACATGCCCTGCACATATCCATGAT 684  
OY 181 GluLeu 182  
DB 685 GAGCTA 690  
RESULT 9  
ID ABR09772 standard; cDNA; 814 BP.  
XX ABR09772:  
AC 14-MAR-2002 (first entry)  
DE Human ovarian tumour protein encoding cDNA #305.  
XX  
XX Human ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;  
KW gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.  
OS Homo sapiens.  
XX  
XX WO200190154-A2.  
XX  
XX 29-NOV-2001.  
XX  
XX 23-MAY-2001; 2001WO-US16895.  
XX  
XX 24-MAY-2000; 2000US-207107P.  
XX 13-JUN-2000; 2000US-211457P.  
XX 21-JUN-2000; 2000US-213673P.  
XX 03-AUG-2000; 2000US-223288P.  
XX 01-MAR-2001; 2001US-272790P.  
XX  
XX (CORI-) CORIAX CORP.  
XX  
XX Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MJ;  
PI Algate PA, Fling SP, Mannion J, Benson DR, Carter D;  
XX  
XX WPI: 2002-097641/13.  
XX  
XX New isolated polynucleotide encoding polypeptide comprising portion of  
PT ovarian tumour protein, useful for detection, diagnosis and therapy of  
PT human ovarian cancer -  
XX  
XX  
PS Claim 1; Page 253-254; 285bp; English.  
XX  
XX The invention relates to an isolated polynucleotide encoding a  
CC polypeptide comprising a portion of an ovarian tumour protein. The  
CC sequences of the invention are useful for stimulating an immune response  
CC and for treating ovarian cancer in a patient. An antigen presenting cell  
CC that expresses the sequences is useful for treating ovarian cancer by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells  
CC can then be proliferated and administered to the patient to inhibit the  
CC development of cancer. The DNA sequences are useful as probes or primers  
CC for nucleic acid hybridisation, to direct expression of a polypeptide in  
CC appropriate host cells. Detecting the presence of a cancer in a patient  
CC involves obtaining a biological sample from the patient, contacting the  
CC biological sample with an agent that binds to the protein, detecting the  
CC amount of protein that binds to the agent, comparing the amount of  
CC protein to a predetermined cut-off value and determining the presence of  
CC cancer. Sequences ABR09464-ABR09802 represent PCR primers and cDNA  
CC molecules encoding ovarian tumour proteins of the invention.

XX  
SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;  
Alignment Scores:  
Pred. No.: 1.55e-99 Length: 814  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 24 Gaps: 0  
US-10-082-502-19 (1-182) x ABR09772 (1-814)  
OY 1 MetLysGlyTyrGlyTyrPheAlaLeuLeuGluGlyValLeuLeuGlyThrAlaTrrAla 20  
DB 145 ATGAAAGCTGGGGTGTGGCTGGCCCTGCTTCTGGGGGCGCTGCGAAGCCCTGGGCT 204  
OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrr 40  
DB 205 CGAGAGACCAAGATCTCCACTGTGGACATGCAAGCGCTGTGGTGAATCAATG 264  
OY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
DB 265 GAAATTCGCCAGGTGAGACCCCAAGAACCATTCAGATGGGATCTTTCGGATCAATCCA 324  
OY 61 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
DB 325 GATGCGACCGACATGCTAGTGTGAGGTGCTTATGCCCTGAGAGCCGACCTCAGAG 384  
OY 81 LeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100  
DB 385 CTGCTGAGAGAAATVATGACCGGATGGAAGATVGGGAAACATGATGATCTTCCACC 444  
OY 101 HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 120  
DB 445 CATGCCAAGAACTACGTACGTGTAGTGGCCGAGATGAGAAATCCAGTAATGACCTA 504  
OY 121 GlnGlyIleAArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
DB 505 CAAAGCATCCGAAATCGACTAGATATAGCGCACCCCTCAAGTTGCGGTGAGAGCAT 564  
OY 141 ValGluGluTyrGlnAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
DB 565 GTGAGAGAAATACGAGATGAACCTCATTAATCTTTCCCGAGAGCTGACAAATGTTAA 624  
OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
DB 625 GACAAACTTTCAGATACGAAACAGATCTTTGTGACATGCCCTGCACATATCCATGAT 684  
OY 181 GluLeu 182  
DB 685 GAGCTA 690  
RESULT 10  
ID AAX22112 standard; DNA; 1180 BP.  
XX AAX22112:  
AC 18-MAY-1999 (first entry)  
DE 18-MAY-1999 (first entry)  
XX  
XX Human secreted protein gene 2 clone H2MBB56.  
DE  
DE Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
KW tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
KW developmental abnormality; foetal deficiency; Alzheimer's disease;  
KW cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
KW immune deficiency disease; respiratory disorder; arthritis; skeletal;  
KW haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
KW cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.  
XX  
XX Homo sapiens.



PR 17-MAY-2000: 2000MO-US13705.  
 PR 22-MAY-2000: 2000MO-US14042.  
 PR 30-MAY-2000: 2000MO-US14941.  
 PR 02-JUN-2000: 2000MO-US15264.  
 PR 10-JUN-2000: 2000MO-US30873.  
 XX  
 PA (GENE) GENENTECH INC.  
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 DR P-PSDB: AAU12258.  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 PS  
 PS Claim 3: Fig 173: 813p: English.  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peritoneal blood  
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 CC  
 XX  
 SQ Sequence 1210 BP: 261 A; 316 C; 379 G; 254 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 2,63e-99 Length: 1210  
 Score: 938.00 Matches: 177  
 Percent Similarity: 98.35% Conservative: 2  
 Best Local Similarity: 97.25% Mismatches: 3  
 Query Match: 97.81% Indels: 0  
 DB: Gaps: 0  
 US-10-082-502-19 (1-182) x AAS21330 (1-1210)  
 OY 1 MetLysGlyTyrGlyTyrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrrAla 20  
 DB 543 ATGAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 602  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyValaCysArgAlaLeuValaAspGluLeuGlnTrr 40  
 DB 603 CGGAGGAGCCGAGATCTCCATGTGAGCATGCAAGGCTCTGTGGATGAACTGAAATGG 662  
 OY 41 GluIleAlaArgValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 663 GAAATGGCCCGAGTGGAGCCCAAGAGACATTCAGATGGGATCTTCCGATCATTCGA 722  
 OY 61 AspGlySerGlnSerValaGluValaProTyrAlaArgSerGluAlaHisLeuThrGlu 80  
 DB 723 GATGGCAGCCAGTCAGTGGTGGAGTGGCTATATGCCCTCAGAGGCCACCATCAGAG 782  
 OY 81 LeuGluGluValaCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100

DB 783 CTGCTGAGAGAGATATGTCACCGCATGAGAGATATGGGAAACAGATTCCTTCACCC 842  
 OY 101 HisArgLysAsnTyrValArgValaValSerArgAsnGlyGlySerSerGluLeuAspLeu 120  
 DB 843 CATGCCAAGAACTACGTACGTGTACTGGCCGAGATGAGAAATCCAGTAATGGACCTTA 902  
 OY 121 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 903 CAAGCAGATCCGAAATCGATCATGATATATAGCCGACCCCTCAAGTTTGCGTGCAGACATT 962  
 OY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 DB 963 GTGGAGAAATACGAGATGAATCAATTAATTTCTTCCCGAGAGCTGACAAATGTAA 1022  
 OY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 1023 GACAAACCTTGACGTACGAGCAAGATCTTTGTGACATGCCCTGCACATATCGCATGAT 1082  
 OY 181 GluLeu 182  
 DB 1083 GAGCTA 1088  
 RESULT 12  
 AA97884  
 ID AA97884 standard; cDNA: 714 BP.  
 XX  
 AC AA97884;  
 XX  
 DT 23-SEP-1999 (first entry)  
 XX  
 DE Human secreted protein encoding cDNA #72.  
 XX  
 KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
 KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN W09925825-A2.  
 PD  
 PD 27-MAY-1999.  
 XX  
 PF 13-NOV-1998; 98WO-IB01862.  
 XX  
 PR 04-SEP-1998; 98US-0099273.  
 PR 13-NOV-1997; 97US-0066677.  
 PR 17-DEC-1997; 97US-0069957.  
 PR 09-FEB-1998; 98US-0074121.  
 PR 13-APR-1998; 98US-0081563.  
 PR 10-AUG-1998; 98US-0096116.  
 XX  
 PA (GENE) GENSET.  
 PA  
 PI Bougueleret L, Duclert A, Dumas Mline Edwards J;  
 PI WPI: 1999-347472/29.  
 DR P-PSDB: AAY36200.  
 PT  
 PT Extended cDNAs encoding secreted proteins  
 XX  
 PS Claim 1; Page 272; 307p; English.  
 XX  
 CC AAX97813-X97906 represent extended cDNA's which encode novel human  
 CC secreted proteins (see AAY36129-Y36222) and which have cytosolic,  
 CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
 CC express secreted proteins or parts of them or to obtain antibodies  
 CC capable of binding to the secreted proteins. They may also be used in  
 CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC Uses also include design of expression vectors and secretion vectors.  
 CC  
 XX  
 SQ Sequence 714 BP: 203 A; 158 C; 202 G; 151 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 8.47e-99 Length: 714



Score: 931.00  
 Percent Similarity: 97.808  
 Best Local Similarity: 96.708  
 Query Match: 97.08%  
 DB: 20  
 Matches: 176  
 Conservative: 2  
 Mismatches: 4  
 Indels: 0  
 Gaps: 0  
 US-10-082-502-19 (1-182) x AAK97884 (1-714)

QY 1 MetLeuGlyTTPGlyTTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB 33 ATGMAAGGCTGGGGTGGCTGGCTGGCTGGGGGCTCTGGGAACCGCTGGGCT 92  
 QY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 40  
 DB 93 CGAGAGAGCCAGCATCTCCACTGTGTGAGCATGAGGCTGTGTGTGAGTACACTACAAATGG 152  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB 153 GAAATTTGCCAGCTGAGCCCAAGAACCATTCAGATGGGATCTTCCGATCATCTCA 212  
 QY 61 AspGlySerGlnSerValAlaGluValProTyrAlaAspSerGluAlaHisLeuThrglu 80  
 DB 213 GATGACAGCCAGTCACTGTGTGAGGCTGTGTATGCCCTCAGAGGCCACCTCAGAGAG 272  
 QY 81 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 100  
 DB 273 CTGCTGAGAGAGATGTGTGACCGGATGAGAGATGTGGGAAACATGTGATCTTCCACC 332  
 QY 101 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 120  
 DB 333 CATGCCAAGAACTACGTACGTGTGTGAGGCTGTGTGTGAGTACGAACTGAGAACTGGACCTA 392  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB 393 CAAGGCATCCGAATCGACTCGATTTAGCGGACCTCAAGTTGGCTGTGGAGCATTT 452  
 QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160  
 DB 453 GTGAGAGAAATGAGAGATGAACTCATTTCTTTCCGAGAGGCTGACAAATGTTAA 512  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB 513 GACAAACTTTTGCACTAAGCAACGACATCTTTGTGACCATGCTCGCACATATCCATGAT 572  
 QY 181 GluLeu 182  
 DB 573 GAGCTA 578

RESULT 13  
 ABK36007  
 ID ABK36007 standard; cDNA: 657 BP.  
 XX  
 AC ABK36007;  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE cDNA sequence #398 encoding novel human secreted protein.  
 XX  
 KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
 KW Immune deficiency disorder; blood disorder; inflammatory disorder;  
 KW Infectious disorder; allergic condition; neurodegenerative disorder;  
 KW Liver fibrosis; coagulation disorder; gene therapy; antitubercular;  
 KW tumor; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200177289-A2.  
 XX  
 PD 18-OCT-2001.  
 XX  
 PF 29-MAR-2001; 2001WO-US10232.  
 XX  
 PR 06-APR-2000; 2000US-195605P.

XX (GENY ) GENETICS INST INC.  
 PA Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
 PT Clark HF, Fechel K, Howes SH, Resnick RJ, Gultokta K, Graham JR;  
 DR WPI: 2002-179322/23.  
 XX

Six hundred and twenty three polynucleotides derived from a variety of human tissue sources which encode secreted proteins, useful for treating immune deficiencies and disorders such as autoimmune disorders  
 -  
 Claim 1; Page 297; 393pp; English.

CC The present invention relates to the isolation of novel cDNA sequences  
 CC which encode human secreted proteins. The cDNA sequences have been  
 CC derived from a variety of human tissues. The invention also provides  
 CC a method for producing proteins from these polynucleotide sequences.  
 CC The proteins are useful for identifying compounds that modulate their  
 CC activity and production. The sequences of the invention are  
 CC useful for treating diseases such as hyperproliferative disorders  
 CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
 CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
 CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
 CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
 CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
 CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
 CC haemophilia), and tumours. The polynucleotide sequences of the  
 CC invention are also useful in gene therapy. ABK3610-ABK36232 represent  
 CC the cDNA sequences of the invention that encode for novel human  
 CC secreted proteins.  
 XX

SO Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;

Alignment Scores:  
 Pred. No.: 8,916-95  
 Score: 896.00 Length: 657  
 Percent Similarity: 97.19% Matches: 170  
 Best Local Similarity: 95.51% Mismatches: 3  
 Query Match: 93.43% Indels: 5  
 DB: 24 Gaps: 0

US-10-082-502-19 (1-182) x ABK36007 (1-657)

QY 5 GlyTTPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAlaArgSerGln 24  
 DB 1 GGTGGCTGGCTCTCTCTGGGGGCTGTGGGAACCGCTGGGCTGGAGGCCAG 60  
 QY 25 AspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrpGluIleAlaArg 44  
 DB 61 GATCTCCAGCTGTGAGCATGAGGCTGTGTGATGATGAAATGGAAATGGCCAC 120  
 QY 45 ValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGln 64  
 DB 121 GTGGACCCCAAGAACCATCATGTGGATCTTCCGAGTCAATCCAGATGGAGCCAG 180  
 QY 65 SerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGluLeuGluGlu 84  
 DB 181 TCAGTGGTGGAGGTCCTTATGCCGCTCAGAGGCCACATCAGAGCTGTGGAGAG 240  
 QY 85 ValCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThrHisArgLysAsn 104  
 DB 241 ATATGTGACCGAGTGAAGAGATATGGGAACAGATATCTTCCACCATCCAGAAC 300  
 QY 105 TyrValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeuGlnGlyIleArg 124  
 DB 301 TACGTACGTGTAGTGGCCGGAATGAGAAATCCAGTCACTGAGACCTACAAAGCATCCGA 360  
 QY 125 IleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyr 144  
 DB 361 ATCGACTCAGATATTTACGCGCACCTCAAGTTTCGCTGTGAGAGCATGTGGAGAAATAC 420

OY 145 GluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAspLysLeuCys 164  
|||||  
Db 421 GAGGATGACCATTCGATTCCTTTCCGACAGCTGACAAATGTTAAAGACAAACTTGGC 480  
OY 165 SerLysAspGlnThrAspLeuCysAspHisAlaLeuHisArgSerHisAspGluLeu 182  
|||||  
Db 481 AGTAAAGCAACGATCTTTGTGACCATGCGCTGACATATGCGATGATGAGCTA 534  
RESULT 14  
ID AAX97837 standard; cDNA: 832 BP.  
XX AAX97837;  
AC AAX97837;  
XX 23-SEP-1999 (first entry)  
XX Human secreted protein encoding cDNA #25.  
DE  
XX Secreted protein: human; cytosolic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX Homo sapiens.  
OS  
XX MO9925825-A2.  
PN  
XX 27-MAY-1999.  
PD  
XX 13-NOV-1998; 98WO-1B01862.  
XX  
XX 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
XX (GSEST ) GENSET.  
PA  
XX Bougueleret L, Duclert A, Dumas Milne Edwards J;  
PI  
XX WPI: 1999-347472/29.  
DR P-PSDB: AAY36153.  
XX  
XX Extended cDNAs encoding secreted proteins  
PT  
XX Example 28; Page 194-195; 307pp; English.  
XX  
XX AAX97813-X97906 represent extended cDNA's which encode novel human  
CC secreted proteins (see AAY36129-X36222) and which have cytosolic,  
CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
CC express secreted proteins or parts of them or to obtain antibodies  
CC capable of binding to the secreted proteins. They may also be used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC Uses also include design of expression vectors and secretion vectors.  
XX  
SO Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other:  
Alignment Scores:  
Pred. No.: 7.95e-94 Length: 832  
Score: 889.00 Matches: 175  
Percent Similarity: 96.22% Conservative: 3  
Best Local Similarity: 94.59% Mismatches: 4  
Query Match: 92.70% Indels: 3  
DB: 20 Gaps: 0  
US-10-082-502-19 (1-182) x AAX97837 (1-832)  
OY 1 MetLysGlyTyrGlyTyrPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTyrPala 20  
|||||  
Db 148 ATGAAGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 207  
OY 21 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGlu-LeuGluTtr 40  
|||||

Db 208 CGGAGAGCCAGGATCTCAGCTGTGAGACATGACGGCTGTGGAGAAACAGATG 267  
OY 40 pGluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAspPr 60  
|||||  
Db 268 GGAATTCGCCAGGAGGAGCCCAAGAACATTCAGATGGATCTTCCGATCAATCC 327  
OY 60 oAspLysSerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGl 80  
|||||  
Db 328 AGATGGAGCCAGTCAGTGTGGAGCTCTTATGCCCGCTCAGAGGCCACCTCACGA 387  
OY 80 uLeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerTh 100  
|||||  
Db 388 GCTGCTGGAGGATATGTGACCGGATGAGAGATATGGAGAACGATGTGATCCTCCAC 447  
OY 100 rHisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLe 120  
|||||  
Db 448 CCATCGCAAGAACATACGATGCTGATGCTGGCCGGAATGGAATTCAGTACAGTGCACCT 507  
OY 120 uGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLys-PheAlaCysGluSerI 140  
|||||  
Db 508 ACAAGGCATCCGAATCGACTCAGATATAGCGGCACCTCAAGBTTCGTGGAGCA 567  
OY 140 lValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValL 160  
|||||  
Db 568 TTGTGGAGGATACGAGATGACATGATTCCTTTCCGAGAGGCTGACATGTTA 627  
OY 160 YAspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSer-His 179  
|||||  
Db 628 AAGCAAACTTGGCAGTAAAGGAAACAGATCTTTGTGACCAATGCCCTGCAATTCGGCAT 687  
OY 180 AspGluLeu 182  
|||||  
Db 688 GATGAGCTA 696  
RESULT 15  
ID AAA43525 standard; cDNA: 592 BP.  
XX AAA43525;  
XX  
XX AAA43525;  
AC  
XX  
XX 21-AUG-2000 (first entry)  
DE  
XX Mouse secreted expressed sequence tag SEQ ID NO:100.  
XX  
XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;  
KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;  
KW antiviral; antidiabetic; antiallergic; antiparasitic; antiparkinsonian;  
KW antitumor; osteoprotective; neuroprotective; nootropic; antiparasitic;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW hemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
OS Mus musculus.  
XX  
XX WO200021991-A1.  
PN  
XX 20-APR-2000.  
PD  
XX  
XX 15-OCT-1999; 99WO-US24206.  
PF  
XX  
XX 15-OCT-1998; 98US-0104436.  
PR  
XX (GENY ) GENETICS INST INC.  
PA  
XX  
XX Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;  
PI





GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1725.3 Seconds

(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-19

Perfect score: 959

Sequence: 1 MKGGMALILGLVLTGTAMA.....LCSKRDPLCHALHRSDEL 182

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xip  
-O=/cg2\_1/USPTO.spool/US10082502/runat.24032003.135059.6547/app-query.fasta.1.1308  
-DB=Genbm1 -GFM=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US10082502.ecgn.1.1.4758.grnat.24032003.135059.6547 -NCP=6 -TCP=3  
-NO\_XLPXY -NO\_MAMP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Genbm1:\*  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sy:\*  
28: em\_un:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	100.0	798	10 BC008261	BC008261 Mus muscu
2	959	100.0	1017	10 AF186115	AF186115 Mus muscu
3	959	100.0	158198	2 AC012013	AC012013 Homo sapi
4	938	97.8	814	6 AX440456	AX440456 Sequence
5	938	97.8	814	9 AB015631	AB015631 Homo sapi
6	938	97.8	1080	9 AY032624	AY032624 Homo sapi
7	938	97.8	1210	6 AX464040	AX464040 Homo sapi
8	474.5	49.5	155023	2 AC025574	AC025574 Homo sapi
9	472	49.2	182350	2 AC073896	AC073896 Homo sapi
10	472	49.2	184762	2 AC073896	AC073896 Homo sapi
11	451.5	47.1	207024	10 AC030500	AC030500 Homo sapi
12	429	44.7	101666	2 AC090489	AC090489 Genomic s
13	361	37.6	649	9 AF186113	AF186113 Homo sapi
14	361	37.6	824	9 BC001027	BC001027 Homo sapi
15	336.5	35.1	69753	6 AC103156	AC103156 Rattus no
16	262	27.3	564	6 AX440375	AX440375 Sequence
17	248	25.9	2378	9 AK094445	AK094445 Homo sapi
18	212.5	22.2	996	3 AY094698	AY094698 Drosophi
19	193.5	20.2	40824	3 U13070	U13070 Caenorhabdi
20	159	16.6	785	8 AY088187	AY088187 Arabidops
21	156.5	16.3	622	5 AB055671	AB055671 Dario fer
22	145	15.1	175963	3 AC009379	AC009379 Drosophi
23	145	15.1	279530	3 AE003518	AE003518 Drosophi
24	142.5	14.9	1286	3 AY060987	AY060987 Drosophi
25	140	14.6	55359	2 AC012854	AC012854 Drosophi
26	140	14.6	172904	3 AC007414	AC007414 Drosophi
27	140	14.6	275390	3 AE003831	AE003831 Drosophi
28	136.5	14.2	744	6 AR142811	AR142811 Sequence
29	136.5	14.2	746	6 AX136616	AX136616 Sequence
30	136.5	14.2	747	6 BD006701	BD006701 Novel pol
31	136.5	14.2	1413	9 BC015903	BC015903 Homo sapi
32	136.5	14.2	1419	6 AX136423	AX136423 Sequence
33	136.5	14.2	1470	6 AX376412	AX376412 Sequence
34	136.5	14.2	1470	6 AX454716	AX454716 Sequence
35	136.5	14.2	1470	6 AX454716	AX454716 Sequence
36	136.5	14.2	1512	6 AR142810	AR142810 Sequence
37	136.5	14.2	1547	6 BC033339	BC033339 Sequence
38	133	13.9	161566	9 AC008060	AC008060 Homo sapi
39	127.5	13.3	76187	2 AC015219	AC015219 Homo sapi
40	123	12.8	168418	2 AC009403	AC009403 Homo sapi
41	116	12.1	1260	6 AX380553	AX380553 Sequence
42	116	12.1	181896	6 AC114126	AC114126 Rattus no
43	116	12.1	201649	2 AC120813	AC120813 Rattus no
44	115.5	12.0	201649	2 AC120813	AC120813 Rattus no
45	113	11.8	210631	2 AC129184	AC129184 Mus muscu

RESULT 1

## ALIGNMENTS

BC008261 LOCUS BC008261 798 bp mRNA linear ROD 07-AUG-2002  
 DEFINITION Mus musculus, transmembrane protein 4, clone MGC:6853  
 IMAGE:2650612, mRNA, complete cds.  
 ACCESSION BC008261  
 VERSION BC008261.1 GI:14198400  
 KEYWORDS MGC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 798)  
 REFERENCE Direct Submission  
 TITLE Submitted (22-May-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 JOURNAL NIH-MGC Project URL: http://mgc.ncl.nih.gov  
 REMARK Contact: MGC help desk  
 COMMENT Email: gcapbs-rt@mail.nih.gov  
 Tissue Procurement: Lotmar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: IRAC Plate: 5 Row: b Column: 23  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA g1: 9903606.  
 FEATURES  
 source  
 1. 798  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="129,C57BL/6J,FVB/N"  
 /clone="MGC:6853 IMAGE:2650612"  
 /tissue\_type="Mammary tumor. Brca1-/-; MTV-Cre model. 10  
 months old, gross tissue."  
 /clone\_id="NCI\_CGAP\_Mam3"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 115..663  
 /codon\_start=-1  
 /product="transmembrane protein 4"  
 /protein\_id="AAH08261.1"  
 /db\_xref="GI:14198401"  
 /db\_xref="locusid:56530"  
 /translation="MKGWGLALLGLVLTSTNARRSODLHCACRLVDELEIKR  
 VDKRTTQSGFRINPDGSSVVEVPARSQAHTLEEVCDMKRKGQIDPSTIR  
 KNYRVVSRGSESELDLOGIRIDISGTLKFACEISIVEYEDLEIEFSSREADNVK  
 DKLSKRTDLCDAHLHRSDEL"  
 BASE COUNT 227 a 182 c 228 g 161 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.29e-88 Length: 798  
 Score: 959.00 Matches: 182  
 Percent Similarity: 100.008 Conservative: 0  
 Best Local Similarity: 100.008 Mismatches: 0  
 Query Match: 100.008 Indels: 0  
 DB: 10 Gaps: 0  
 US-10-082-502-19 (1-182) x BC008261 (1-798)  
 QY 1 MetylsGlyTPpGlyTPpleuAlaLeuLeuGlyValLeuLeuGlyThAlaTrpAla 20

|||||  
 Db 115 ATGAAGGCTGGGGTGTGGCTACCTACTTGGGGGCTCGTGGGAACGCCGGGCT 174  
 QY 21 ArgArSerGlnAspLeuHiscysGlyAlaCysAArgAlaLeuValAspGluTrp 40  
 Db 175 CGAAGGACCAAGATCTACACTGTGAGCTTCAGAGGCTCTGGTGAATTAAGTGG 234  
 QY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 Db 235 GAAATGGCCCGCTGGAGCCCAAGAACATTCAGATGGGATCCTTCGATCAATCCA 294  
 QY 61 AspGlySerGlnSerValValGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 80  
 Db 295 GATGGCACCGACGTCAGTGTGGAGGTACTTATGCCCTCAGAGGCCACCTCCACAG 354  
 QY 81 LeuLeuGluValAlaCysAspArgMetLysGlyTrpGlyGluGlnIleAspProSerThr 100  
 Db 355 TTGCTTGAAGAGAGTGTGACCGAATGAAGAGTACGGGGAACAAATGACCTCTAC 414  
 QY 101 HisArgLysAsnTrpValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 120  
 Db 415 CACCCCAAGAACTAGATGATCGCGTCTGAGCCGGAATGGAATCCAGTAAGTACCTTA 474  
 QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 Db 475 CAGGCGATCCGAATGACTCAGATATCAGCGCACCCCTCAAGTTGGCTGTAGACATT 534  
 QY 141 ValGluGluTrpGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 Db 535 GTGGAAGATAGAGAGATGAGACTTATCGAATCTCTCCAGAGAGCTGACACACTTAA 594  
 QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 Db 595 GACAAACTTTCAGATGACGCGACAGATCTATGTGCCATGCCCTCAGACAGATTCAGAT 654  
 QY 181 GluLeu 182  
 Db 655 GAGCTA 660  
 RESULT 2  
 AF186115 1017 bp mRNA linear ROD 13-JAN-2000  
 LOCUS AF186115  
 DEFINITION Mus musculus putative secreted protein ZS199 (Zs199) mRNA, complete  
 cds.  
 ACCESSION AF186115  
 VERSION AF186115.1 GI:6014635  
 KEYWORDS  
 SOURCE  
 ORGANISM Mus musculus.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1017)  
 REFERENCE Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Mus musculus putative secreted protein  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1017)  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics,  
 Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
 FEATURES  
 source  
 1. 1017  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="dbEST:AA049839"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:476061"  
 1..1017  
 /gene="Zs199"  
 343..891  
 /gene="Zs199"  
 /codon\_start=-1

/product="putative secreted protein ZS19"  
 /protein\_id="AA01433.1"  
 /db\_xref="gi:6014636"  
 /translation="MKGMWLLALLGLTGMARRSODLHCACRAVDELEMEIAR  
 VPKRTIOMGSRPRINDGOSVSEVPRSEAHNTELEEEYCDNRKEGEIDISTHR  
 KNYVAVRNGSESELDIGRIRISDISIGTLKFCESTVEEYEDLELFFESREADNVK  
 DKICSRKRDLDHAIHRSHDEL"

BASE COUNT 248 a 250 c 315 g 204 t

ORIGIN

Alignment Scores:

Pred. No.:	4.44e-88	Length:	1017
Score:	959.00	Matches:	182
Best Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-082-502-19 (1-182) x AF186115 (1-1017)

OY 1 MetlySGlyTTPGLYTRPLeuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrpAla 20  
 DB ATGAAGGCTGGGGTGGCTAGCCCTACTTTGGGGGCTCTGCTGGAACTGCGTGGCT 402  
 OY 21 ArgArgSerGlnAspLeuHisCysGlyValAlaCysArgAlaLeuValAspGluLeuGluTrp 40  
 DB CGAAGGACCAAGATCTCACTGAGAGCTGCAAGGCTCTGGTGGATGAAATTAAGAGTGG 462  
 OY 41 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 60  
 DB GAAATTCGCCCGCTGACCCCAAGAACCACTTCAGATGGATCTTCGCAATCAATCCA 522  
 OY 61 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 80  
 DB GATGGACCCAGTCACTGTTGAGAGTACTTATGCCCGCTAGAGGCCCACTCAGAGAG 582  
 OY 81 LeuLeuGluGluValCysAspArgMetLysGluTrpGlyGluGluIleAspProSerThr 100  
 DB TTGCTTGGAGAGTGTGTGACCAATGAGAGTACCGGGAACGATTTGACCTTCTTACC 642  
 OY 101 HisArgLysAsnGlyTrpValArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 120  
 DB CACCGCAAGAACTACGATGACGCTGTCGAGCCGGAATGAGAACTCACTGAACTGACTTA 702  
 OY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 140  
 DB CAGGCGCATCCGAATTCAGATTCAGATTCAGGCGCACTTCAAGTTGCTGTGAGAGCATTT 762  
 OY 141 ValGluGluTrpGlyLeuAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160  
 DB GTGGAAGAAATACGAGATGAGCTTATGAACTTCTTCCAGAGAGGCTGACAAAGCTTAA 822  
 OY 161 AspLysLeuGlySerLysArgGlyThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180  
 DB GACAAACTTTCGATTAAGCGGACAGATTAATGACCAATGCTCCCTGACAGATTCACAGAT 882  
 OY 181 GluLeu 182  
 DB 883 GAGCTA 888

RESULT 3  
 AC012013 158198 bp DNA linear HTG 17-JUL-2001  
 LOCUS  
 DEFINITION Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT  
 AC012013  
 AC012013 18 GI:14717292  
 VERSION  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFTN.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 158198)

## AUTHORS

Muzny D.M., Adams C., Adio-Oduola B., Ali-Usman F.R., Allen C.,  
 Alsbrooks S.L., Amaratunga H.C., Are J.R., Banks T., Barbara J.,  
 Beaton J., Birmage K., Blankenburg K., Bonlin D., Bouck J.,  
 Bowle S., Brileva M., Brown M., Brown M., Bryant N.P., Buha C.,  
 Burck P., Burkett C., Burrell K.L., Byrd N.C., Caron T.F.,  
 Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R.,  
 Chen Z., Chowdhury I., Christopoulos C., Cleveland C.D., Cox C.,  
 Coyle M.D., Dethorne S.R., David R., Davila M.L., Davis C.,  
 Davy-Carroll L., Dederich D.A., Delaney K.R., Delgado O.,  
 Duggan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C.,  
 Elhaj C., Escotto M., Falls T., Ferraguto D., Flagg N., Ford J.,  
 Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T.,  
 Garza N., Gill R., Gorrell J.H., Guvera M., Gunaratne P., Hale S.,  
 Hamilton K., Harris C., Harris R., Hart M., Havlak P., Hayes A.,  
 Hernandez J., Hernandez O., Hodgson A., Hognes M., Holloway C.,  
 Hollins B., Homsl F., Howard S., Huber J., Huliy S., Hume J.,  
 Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S.,  
 Joudah S., Karlsson E., Kelly S., Khan U., King L., Korvah J.,  
 Kovar C., Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C.,  
 Lewis L., Li J., Li Z., Lichtenage O., Lieu C., Liu J., Liu W.,  
 Louieged H., Lozano R.J., Lu X., Lucier A., Lucier R., Luna R.,  
 Ma J., Maheshwari M., Mapua P., Martin R., Martindale A.,  
 Martinez E., Massey E., Mawlin E., McLeod M.P., Meador M.,  
 Mel G., Metzker M., Miner G., Miner Z., Mitchell T., Mohabbat K.,  
 Morgan M., Morris S., Moser M., Neal D., Newton J., Newton N.,  
 Nguyen A., Nguyen N., Nguyen N., Nickerson E., Nwokweto S.,  
 Ogun M., Okunoye G., Oragunye N., Oviedo R., Pace A., Payton B.,  
 Peery J., Perez L., Peters L., Pickens R., Primus E., Pu L.L.,  
 Quiles M., Ren Y., Rivers M., Rojas A., Rojupokan I., Rolfe M.,  
 Ruiz S., Savery G., Scherer S., Scott G., Shen H., Shooshari N.,  
 Sisson I., Sodergren E., Sponale T., Sparks A., Stanley H.,  
 Stone H., Sutton A., Svatek A., Tabor P., Tamerisa A., Tamerisa K.,  
 Tang H., Tansey J., Taylor C., Taylor V., Villalob D., Vinsan R.,  
 Thomas S., Usmani K., Vasquez L., Vera T., Villalob D., Vinsan R.,  
 Wall R., Wang S., Ward-Moore S., Warren R., Washington C.,  
 Watlington S., Williams G., Williams A., Wlezyk R., Wooten S.,  
 Worley K., Wu C., Wu Y.F., Zhou J., Zorilla S., Nelson D.,  
 Weinstein G. and Gibbs R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

## COMMENT

Direct Submission  
 Unpublished  
 2 (bases 1 to 158198)  
 Worley, K.C.  
 Direct Submission  
 Submitted (19-OCT-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 13, 2001 this sequence version replaced gi.14547727.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: HMKZ  
 Center clone name: RP11-764L14  
 ----- Summary Statistics  
 Sequencing vector: M13; I08821  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 161278 bases at least Q40  
 Consensus quality: 170806 bases at least Q30  
 Consensus quality: 174843 bases at least Q20  
 Estimated insert size: 168803; sum-of-contigs estimation  
 Estimated insert size: 163288; agarose-gel estimation  
 Quality coverage: 6x in Q20 bases; agarose-gel estimation  
 Quality coverage: 5.8x in Q20 bases; sum-of-contigs estimation  
 ----- NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is



\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 108229: contig of 108229 bp in length  
\* 108230 108329: gap of unknown length  
\* 108330 150246: contig of 41917 bp in length  
\* 150247 150346: gap of unknown length  
\* 150347 153410: contig of 3064 bp in length  
\* 153411 153510: gap of unknown length  
\* 153511 155739: contig of 2229 bp in length  
\* 155740 155839: gap of unknown length  
\* 155840 158198: contig of 2359 bp in length.

FEATURES  
Source  
1. 158198  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-764L14"

BASE COUNT 46057 a 30922 c 31747 g 49030 t 442 others  
ORIGIN

## Alignment Scores:

Pred. No.: 2,26e-85 Length: 158198  
Score: 959.00 Matches: 182  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-082-502-19 (1-182) x AC012013 (1-158198)

OY 1 Metlysglyttrpglyttrpleualaleuleuleuglyvalleuleuglythralatpala 20  
DB 152111 ATGAAGAGCTGGGGTGGCTAGCCCTACTTTGGGGTCTCGTGGAACTGGGCT 152170  
OY 21 ArgatserglnaapleuHiscysgylvalacyatgatalaleuvalaasgduleuglutr 40  
DB 152171 CGAAGGACCAAGATCTACCTGTGAGCTTGCAAGGCTCTGTGGATGAAATTAAGTGG 152230  
OY 41 Gluilealargvalaasprolylsythrilleglnmetglsyserpheargilleaspro 60  
DB 152231 GAAATGGCCCGCTGGAGCCCAAGAAAGCAATTCAGATGGGATCTTCCGAATCAATCCA 152290  
OY 61 Aspglyserglnservalvalgluvalprotyrallaargserglualahisleuthrglu 80  
DB 152291 GATGGCACCCAGTCAGTGTGTGAGGTACCTTATGCCCGCTCAGAGGCCACCTCCACAGAG 152350  
OY 81 Leuleuglulvalcysaapargmetlysgluluryrgyluglnlileasproserthr 100  
DB 152351 TTGCTTGAGAGGTGTGTGACCAATGAAGAGTACGGGGAACAAATTTGACCTTCTACG 152410  
OY 101 Hisarglysaantyrvalargvalaalsearargannglylusersergluleuaspneu 120  
DB 152411 CACGGCAAGAACTACGTACGGGTGTGAGCCGGAATGAGAAATCCAGAACTAGACTTA 152470  
OY 121 GlnnglylleargylleaspsersapllesergluthrleuylphealaCysgluserlle 140  
DB 152471 CAGGCAATCCCAATTCAGATCATCAGCGGACCCCTCAAGTTTGCCTGTGAGAGCAT 152530  
OY 141 Valglugluluryrgluaspgluleuileglupheseatarglualaaspsnvallys 160  
DB 152531 GTGGAAGAAATCGAGAGATCTTTCGAATTTCTTCCAAAGAGGCTGAAACGTTTAA 152590  
OY 161 AsplylsleucysserlysargtthrAspleucysasphlialaleuHlsargSerHlsasp 180  
DB 152591 GACAAACTTTTCAGATGAAGCGAAGATCTATGTGACCATGCCCTGCACAGATTCACAGAT 152650  
OY 181 Gluleu 182  
DB 152651 GACCTA 152656

RESULT 4  
AX440456  
LOCUS AX440456 814 bp DNA linear PAT 28-JUN-2002  
DEFINITION Sequence 309 from Patent WO0190154.  
ACCESSION AX440456  
VERSION AX440456.1 GI:21665266  
KEYWORDS

## SOURCE

ORGANISM  
human.

REFERENCE  
1 Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H.,  
Lodes, M.D., Algate, P.A., Fling, S.P., Mannion, J., Benson, D.R. and  
Carter, D.

AUTHORS  
1  
Compositions and methods for the therapy and diagnosis of ovarian  
cancer

TITLE  
JOURNAL  
Patent: WO 0190154-A 309 29-NOV-2001;  
CORIXA CORPORATION (US)

FEATURES  
Source  
1. 814  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 210 a 194 c 241 g 169 t  
ORIGIN

## Alignment Scores:

Pred. No.: 4,54e-86 Length: 814  
Score: 938.00 Matches: 177  
Percent Similarity: 98.35% Conservative: 2  
Best Local Similarity: 97.25% Mismatches: 3  
Query Match: 97.81% Indels: 0  
DB: 6 Gaps: 0

US-10-082-502-19 (1-182) x AX440456 (1-814)

OY 1 Metlysglyttrpglyttrpleualaleuleuleuglyvalleuleuglythralatpala 20  
DB 145 ATGAAGAGCTGGGGTGGCTAGCCCTACTTTGGGGTCTCGTGGAACTGGGCT 204  
OY 21 ArgatserglnaapleuHiscysgylvalacyatgatalaleuvalaasgduleuglutr 40  
DB 205 CGAAGGACCAAGATCTACCTGTGAGCTTGCAAGGCTCTGTGGATGAAATTAAGTGG 264  
OY 41 Gluilealargvalaasprolylsythrilleglnmetglsyserpheargilleaspro 60  
DB 265 GAAATGGCCCGCTGGAGCCCAAGAAAGCAATTCAGATGGGATCTTCCGAATCAATCCA 324  
OY 61 Aspglyserglnservalvalgluvalprotyrallaargserglualahisleuthrglu 80  
DB 325 GATGGCACCCAGTCAGTGTGTGAGGTACCTTATGCCCGCTCAGAGGCCACCTCCACAGAG 384  
OY 81 Leuleuglulvalcysaapargmetlysgluluryrgyluglnlileasproserthr 100  
DB 385 CTGCTGAGAGAAATTCAGTACCGGATGAAGAGATGTGGGGAACAAATTTGACCTTCCACC 444  
OY 101 Hisarglysaantyrvalargvalaalsearargannglylusersergluleuaspneu 120  
DB 445 CATGCAAGAACTACGTACGTGTGAGGTGGCGGAATGAGAAATCCAGAACTAGACTTA 504  
OY 121 GlnnglylleargylleaspsersapllesergluthrleuylphealaCysgluserlle 140  
DB 505 CAAAGCAATCCCAATTCAGATCATCAGCGGACCCCTCAAGTTTGCCTGTGAGAGCAT 564  
OY 141 Valglugluluryrgluaspgluleuileglupheseatarglualaaspsnvallys 160  
DB 565 GTGGAAGAAATCGAGAGATCTTTCGAATTTCTTCCAAAGAGGCTGAAACGTTTAA 624  
OY 161 AsplylsleucysserlysargtthrAspleucysasphlialaleuHlsargSerHlsasp 180  
DB 625 GACAAACTTTTCAGATGAAGCGAAGATCTTGTGACCATGCCCTGCACAGATTCACAGAT 684  
OY 181 Gluleu 182

Db 685 GAGCTA 690

|||||

RESULT 5  
AB015631

LOCUS  
AB015631

DEFINITION  
Homo sapiens mRNA for type II membrane protein, complete cds.

ACCESSION  
AB015631

VERSION  
AB015631.1

KEYWORDS  
type II membrane protein.

SOURCE  
Homo sapiens

ORGANISM  
Homo sapiens

REFERENCE  
1 Yokoyama-Kobayashi, M., Yamaguchi, T., Sekine, S. and Kato, S.  
Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank

JOURNAL  
Gene 228 (1-2), 161-167 (1999)

MEDLINE  
99173880

REFERENCE  
2 (bases 1 to 814)

AUTHORS  
Kato, S.

TITLE  
Direct Submission

JOURNAL  
Submitted (16-JUN-1998) Seishi Kato, Research Institute of National Rehabilitation Center for the Disabled, Department of Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan (E-mail: seishi@rehab.go.jp, Tel: 042-995-3100 (ex.2568), Fax: 042-995-3132)

FEATURES  
source

1. 814

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="HP10390"

/tissue\_type="gastric adenocarcinoma"

/clone\_id="PKA1-meta-1"

145. 693

/codon\_start=1

/product="type II membrane protein"

/protein\_id="BA76498.1"

/db\_xref="GI:4586840"

/translation="MKGWGLALLGLLGTAWARRSODLHGCACRALVDELEMEIAO VDPKKTQMGSEFRINPDGQSIVVEPYARSEAHITLLEICDPMKREYGEIDPSTHR KNYRVVGRNGSESELDLOGIRIDSDISGLFKACSIYEYDELEIEFSSRADNVK DKLCSEKRTDCLDHALHSHDEL"

BASE COUNT 210 a 194 c 241 g 169 t

ORIGIN

Alignment Scores:

Pred. No.: 4.54e-86 Length: 814

Score: 938.00 Matches: 177

Percent Similarity: 98.35% Conservative: 2

Best Local Similarity: 97.25% Mismatches: 3

Query Match: 97.81% Indels: 0

Gaps: 0

US-10-082-502-19 (1-182) x AB015631 (1-814)

QY 1 MetTysGlyTyrPglYtrPleuAlaLeuLeuGlyValLeuLeuGlyThrAlaTrrPaLa 20

Db 145 ATGAAGGCGTGGGCTGGCCCTGCTGCTGGGAGCCGCTGGGCT 204

QY 21 ArgArgSerGlnAspLeuHIsCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrrP 40

Db 205 CGGAGAGCCAGCATCTCCACTGTGAGCATGAGGCGCTGTGTGATCACTAGCATG 264

QY 41 GluTlAlaArgValAspProLysTyrThrIleGlnMetGlySerPheArgIleasnPro 60

Db 265 GAATTTGCCAGGTGAGCCCAAGAACCATTCAGATGAGATCTTCCGATCATCA 324

QY 61 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuHrgu 80

|||||

Db 325 GATGGACGACACTCAGTGTGAGCTGCCTTATGCCCCCTAGAGCCCACTCAGAG 384

QY 81 LeuLeuGluGluValCysAspArgMetLysGlyTyrGlyGluGlnIleAspProSerThr 100

Db 385 CTGCTGGAGGAGATATGTGACCGGATGAGAGATGAGGAACGATGATGATCTTCAC 444

QY 101 HisArgLysAsnTyrValArgValSerArgAsnGlyGlySerSerGluLeuAspLeu 120

Db 445 CATGCCAAGAACTACGTACGTGTAGTGGCCGGATGGAGATCGACGAACTGACCTA 504

QY 121 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlySerIle 140

Db 505 CAAGCATCCGAAATCGACTCAGATATTACCGGCACCTCAAGTTTGGCTGAGAGCAT 564

QY 141 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 160

Db 565 GTGAGGAAATCGAGATGTAACATCATGAAATCTTCCGAGAGGCTGACAAATGTTAA 624

QY 161 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180

Db 625 GACAAACTTTCAGATTAGCGAAGCATCTTGTGACATGCCCCCTGCACATATCGCATGAT 684

QY 181 GluLeu 182

Db 685 GAGCTA 690

RESULT 6

LOCUS  
AY032624

DEFINITION  
Homo sapiens saposin-like protein mRNA, complete cds.

ACCESSION  
AY032624

VERSION  
AY032624.1

KEYWORDS  
GI:20196198

SOURCE  
Homo sapiens.

ORGANISM  
Homo sapiens

REFERENCE  
1 (bases 1 to 1080)

AUTHORS  
Bornhauser, B.C., Olsson, P.-A. and Lindholm, D.

TITLE  
NSAP is a novel saposin-like protein that interacts with MIR and stimulates neurite outgrowth

JOURNAL  
Unpublished

REFERENCE  
2 (bases 1 to 1080)

AUTHORS  
Olsson, P.-A. and Lindholm, D.

TITLE  
Direct Submission

JOURNAL  
Submitted (13-APR-2001) Neuroscience, Uppsala University, Husargatan 3, Uppsala 75123, Sweden

FEATURES  
source

1. 1080

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="12"

/map="12q15"

419. 967

/note="NSAP: encodes a type 2 membrane signal and a putative ER retention motif"

/product="saposin-like protein"

/codon\_start=1

/protein\_id="AAK38148.1"

/db\_xref="GI:20196198"

/translation="MKGWGLALLGLLGTAWARRSODLHGCACRALVDELEMEIAO VDPKKTQMGSEFRINPDGQSIVVEPYARSEAHITLLEICDPMKREYGEIDPSTHR KNYRVVGRNGSESELDLOGIRIDSDISGLFKACSIYEYDELEIEFSSRADNVK DKLCSEKRTDCLDHALHSHDEL"

BASE COUNT 238 a 269 c 343 g 230 t

ORIGIN

Alignment Scores:

Pred. No.: 6.44e-86 Length: 1080

Score: 938.00 Matches: 177

Percent Similarity: 98.35% Conservative: 2

Best Local Similarity: 97.25% Mismatches: 3

Query Match: 97.81% Indels: 0

DB: 9 Gaps: 0

US-10-082-502-19 (1-182) x AY032624 (1-1080)

QY 1 MettysgltYrPbglYrPleuAlaleuLeuEnglyValleuLeuglyThraLatrpAla 20  
 |||||  
 Db 419 ATGAAAGGTGGGGTTGGCTGGCCCTGCTTCGCGGGGGCCCTCTGGAAACCGCTGGGCT 478  
 |||||

QY 21 ArgArSergLlAspLeuHlscysglYalAcYArGaLaLeuValAspLlueuglUtrp 40  
 |||||  
 Db 479 CGAGAGGACGACGACTCCACTGTGGAGCATCAGAGGCTCTGTGGATGAACTAAGATGG 538  
 |||||

QY 41 GllleAlaArValAspProlyslYsrTrlleglImetGlySerPhaArgllAsnPro 60  
 |||||  
 Db 539 GAAATTTGCCAGGTGGAGCCCAAGAAAGACATTCAAGATGGATCTTTCCGATCAATCCA 598  
 |||||

QY 61 AspGlySergLlSerYalValAluValProYrAlaArSergLlAlHlslPheThrglu 80  
 |||||  
 Db 599 GATGGAGCGCACTAGTGGTGGAGGTCCTTATGCCCGCTCAGAGCCACCTCACAAG 658  
 |||||

QY 81 LeuLeuGluGluValAcYAspArgmetlysglUtrYrglyGluGlnlLeaSprSerThr 100  
 |||||  
 Db 659 CTGCTGGAGAAATATGTATGACCAGATGAAGAGTAGTATGGGAGACATGATGCTTCCACC 718  
 |||||

QY 101 HlAsArglyAsnTrYValArgValAlaSerArGAsnGlylUserSergLlueuAspLeu 120  
 |||||  
 Db 719 CATCGGAGAACTACGTACGTGTACTGGGCCGGAATGGAGAACTCAAGTAACTGACCTA 778  
 |||||

QY 121 GlnGlylLeaArglLeaSprSerAspLlSerGlyThrLeuYlPheAlAcysglUserIle 140  
 |||||  
 Db 779 CAAGCATCCGAATGACACACATATTTACGGCACCTCTCACTTGGCGGTGAGACATT 838  
 |||||

QY 141 ValGluGluTrYrgLlAspLlLeuIleGluPhePheSerArGlyAlaAspAsnVallys 160  
 |||||  
 Db 839 GTGGAGGATACGAGAGATGAATCACTTGAATCTTTCCGAGAGCGTACAAATGTAA 898  
 |||||

QY 161 AspylLeuCySerSerlyArGhrAspLeuCyAspArHlAlaLeuHlAspSerIAsp 180  
 |||||  
 Db 899 GACAACTTTCAGTAAAGCAACAATCTTTGTGACCAATGCCCTGCACATATCGCATGAT 958  
 |||||

QY 181 GluLeu 182  
 |||||

Db 959 GAGCTA 964

RESULT 7  
 AX464040 1210 bp DNA linear PAT 16-JUL-2002

LOCUS  
 DEFINITION Sequence 173 from Patent WO0140466.  
 ACCESSION AX464040  
 VERSION AX464040.1 GI:21899037  
 KEYWORDS  
 SOURCE  
 ORGANISM human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,  
 Gao,M.Q., Gerlitsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,  
 Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,  
 Wood,W.L. and Zhang,Z.  
 Secreted and transmembrane polypeptides and nucleic acids encoding  
 same  
 Patent: WO 0140466-A 173 07-JUN-2001;  
 Genentech Inc. (US)

JOURNAL  
 FEATURES  
 source Location/Qualifiers  
 1..1210  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 261 a 316 c 379 g 254 t  
 ORIGIN

Alignment Scores:  
 7.41e-86 Length: 1210

Score:	938.00	Matches:	177
Percent Similarity:	98.35%	Conservative:	2
Best Local Similarity:	97.25%	Mismatches:	3
Query Match:	97.81%	Indels:	0
DB:	6	Gaps:	0
US-10-082-502-19 (1-182) x AX464040 (1-1210)			
QY 1	MettysglyTrpGlyTrpLeuAlaLeuLeuGlyValLeuGlyThrAlaTrpAla 20		
DB 543	ATGAAGAGGTGGGGTGGCTGGCTGGCTTCTGGGGGCCCTCTGGAAACGGCTGGGCT 602		
QY 21	ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 40		
DB 603	CGAGAGAGCCAGAACTCCACTGTGGAGATCCAGGGCTCTGTGTGAATCAATCAATGG 662		
QY 41	GluLeuAlaArgValAspProGlySerThrIleGlnMetGlySerPheArgIleAsnPro 60		
DB 663	GAAATTTGCCAGGTGGACCCCAAGAACCAATTCAGATGGGATCTTTCCGGATCAATCCA 722		
QY 61	AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 80		
DB 723	GATGGCAGCACTAGTGGTGGAGGTGGCTTATGGCCCGCTGAGAGGCCACCTCAAGAG 782		
QY 81	LeuLeuGluGluValCysAspArgMetCysGlyTyrGlyGluGlnIleAspProSerThr 100		
DB 783	CTGCTGGAGGAGATATGTGACCCGATGGAAGAGTATGGGGAACAGATTCCTTCACAC 842		
QY 101	HisArgGlyAsnTyrValArgValValSerArgAsnGlyGluSerGlnLeuAspLeu 120		
DB 843	CATGCCAGAACTACGTACGTACTGGGCGCCGGAATGGAGAACTCAAGTGAACCTGAGACCTA 902		
QY 121	GlnGlyIleArgIleAspSerAspIleSerGlyThrIleuValPheAlaCysGluSerIle 140		
DB 903	CNAGGCATCCGAATGAGACTCAGATATTGAGCGGCACCCCTCAATTTGGCGTGAAGCANT 962		
QY 141	ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 160		
DB 963	GTGGAGGAAATGAGAGAACTCATTTGATCTTTCCCGAGAGCGTGCACATGTGTAA 1022		
QY 161	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 180		
DB 1023	GACAAACTTTGCAGTAAAGCAACACATCTTGTGACACATGCCCTGCACATATGCAATGAT 1082		
QY 181	GluLeu 182		
DB 1083	GAGCTA 1088		
RESULT 8			
AC023574/C	155023 bp	DNA	linear
AC023574	155023 bp	DNA	linear
LOCUS	155023 bp	DNA	linear
DEFINITION	Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT		
SEQUENCE	12 unordered pieces.		
AC023574	AC023574.13	GI:20428723	
VERSION	AC023574.13	GI:20428723	
KEYWORDS	HTG; HTGS; PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 155023)		
	Muny D.M., Adams C., Adio-Oduola B., Ali-ossan F.R., Allen C.,		
	Albrooks S.L., Amaralunge H.C., Are J.R., Ayale M., Banks T.,		
	Barbata J., Benton J., Blmage K., Blankenburg K., Bonnin D.,		
	Bouch J., Bowle S., Bileva M., Brown E., Brown M., Bryant N.P.,		
	Bunay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,		
	Carroll T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,		
	Chen G., Chen R., Chen Z., Chowdhury I., Christopoulos C.,		
	Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,		
	Dayala M.L., Davis C., Davy-carroll L., Deedrich D.A.,		
	Delaney K.R., Delgado O., Denn A.L., Ding Y., Dlin H.H.,		
	Donthwaite C.J., Draper H., Dugan-Rocha S., Durbin K.J.,		
	Earnhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,		



OY 168 ----- 168  
 Db 55640 CTGTTCTAGTACTCTGTGGAACCTGACATTCCTTACTACCCATACCCCATCCCA 55581  
 OY 168 ----- 168  
 Db 55580 GAAACCTCTGATTAGTCAAGAGCTGCTTCCATTTCCCTTGCTTGAAGGACTGGG 55521  
 OY 168 ----- 168  
 Db 55520 GTGTTCTCTCACATTTTCTGTACGCTTCAGAAATTTGTGCCCTTGACCTGAGATACA 55461  
 OY 168 ----- 168  
 Db 55460 GGTGAGAGACAGTGGGATGTTTATCATTTATAGCATATATTCGACAGATGTGCT 55401  
 OY 168 ----- 168  
 Db 55400 TCGTCTGCTTTGTGTACAGCTTGTGCGGAATTTTCTGATCTGACATATACGTA 55341  
 OY 169 ----- 175  
 Db 55340 TGAGATGTCATATTCCTTAAATTTGTCGATTTGCGATCTTGTGACATGCTGCTG 55281  
 OY 176 H1SAGSERRHISAPGLUEN 182  
 Db 55280 CACATATGTCATGATGAGCTA 55260

RESULT 9  
 AC073896/c 182350 bp DNA linear HTG 24-AUG-2002  
 LOCUS Homo sapiens chromosome 12 clone RP11-977G19, WORKING DRAFT  
 DEFINITION  
 AC073896  
 AC073896.37 GI:22450359  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULMOP; HTGS\_ACTIVEPIN.  
 ORGANISM human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 182350)

REFERENCE  
 AUTHORS  
 Muzny, D.M., Adams, C., Adio-Oduola, B., All-oshman, F.R., Allen, C.,  
 Alstbrooks, S.L., Amertunge, H.C., Arc, J.R., Ayale, M., Banks, T.,  
 Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowler, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Deedrich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dihn, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franze, P.,  
 Gabisi, A., Geo, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Homsi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kralovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Licharge, O., Lieu, C., Liu, C., Liu, W., Louisgied, H.,  
 Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Matlindale, A., Martinez, E.,  
 Messer, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, A., Morris, S.,  
 Moser, M., Neel, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenwo, S., Ogund, M., Okunodu, G.,  
 Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,

Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Sytek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, O.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleccyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, D., Zorrilla, S., Nelson, D.,  
 Weinstein, G. and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 182350)  
 Worley, K.C.  
 Direct Submission  
 Submitted (01-JUL-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 182350)  
 Worley, K.C.  
 Direct Submission  
 Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Aug 23, 2002 this sequence version replaced gi:22296903.

## COMMENT

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Drafting Center Code: BCM  
 Contact: hgsc-help@bcm.tmc.edu

Project Information  
 Center project name: HBW1  
 Center clone name: RP11-977G19

Summary Statistics

Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 15% of reads  
 Chemistry: Dye-terminator Big Dye; 85% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 179279 bases at least 40  
 Consensus quality: 180581 bases at least Q30  
 Consensus quality: 181073 bases at least Q20  
 Estimated insert size: 17425; sum-of-contigs estimation  
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 2893: contig of 2893 bp in length  
 2894 2993: gap of unknown length  
 2994 6686: contig of 3693 bp in length  
 6687 6786: gap of unknown length  
 6787 17251: contig of 10465 bp in length  
 17252 17352: gap of unknown length  
 17352 85894: contig of 68543 bp in length  
 85895 85994: gap of unknown length  
 85995 182350: contig of 96356 bp in length.

## FEATURES

## SOURCE

1. 182350  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-977G19"

BASE COUNT 46453 a 45416 c 43855 g 46174 t 452 others

## ORIGIN

Alignment Scores:

Pred. No.: 6,35e-36 Length: 182350  
 Score: 472.00 Matches: 98  
 Percent Similarity: 65.13% Conservative: 1  
 Best Local Similarity: 64.47% Mismatches: 1  
 Query Match: 49.22% Indels: 52  
 DB: 2 Gaps: 1

US-10-082-502-19 (1-182) x AC073896 (1-182350)

69 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluGluValCysAspArg 88  
 Db 172027 GTCGCTTATGCCCCGCTCAGAGCCGCTCAGAGCTGCTGAGAGATATGTGACCG 171968  
 Oy 89 MetTyrGluTyrGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 108  
 Db 171967 ATGAAGAGATATGAGGAGAGATATGATCTTCACCCATGCGAAGATACGATAGCTGTA 171908  
 Oy 109 ValSerArgAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 128  
 Db 171907 GTGGGCGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 171848  
 Oy 129 IleSerGlyThrLeuTyrAspPheAla 136  
 Db 171847 ATTAGCGGACCCCTCAAGTTGGGGGAGCTATGGAATCGTAGCTGTTCTTGACAT 171788  
 Oy 136 136  
 Db 171787 TAAGGGGTGTGGAGAGACCAATTAGAAAGTATAGCTGATATCTTCTCTCTTGG 171728  
 Oy 136 136  
 Db 171727 AAGTGAAGCAAGAGCCCTCTTATGCTCCCTGTCTGACCACTTCTCTGTTGGTCAG 171668  
 Oy 137 CysGluSerIleValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 156  
 Db 171667 TGTGAGAGCATTTGTGAGAGATATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 171608  
 Oy 157 AspAsnValIleAspIleAspIleAspIleAspIleAspIleAspIleAspIleAspIle 168  
 Db 171607 GACAAATGTTAAAGCAAACTTGTGACAGTAAGCA 171572  
 RESULT 10  
 LOCUS AC023500 184762 bp DNA linear HTG 24-AUG-2002  
 DEFINITION Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
 AC023500  
 AC023500  
 AC023500 32 GI:22095070  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 184762)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barbarta,J., Benton,J., Bimaye,K., Blankenburg,K., Bryant,N.P.,  
 Burch,J., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinu,H.H.,  
 Douhwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Hawlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgeson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Hombl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loubege,H.,  
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapue,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Mohabbat,K., Morgan,M., Morris,S.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunolu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rivers,M., Rojas,A., RojudoXan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,  
 Sodergren,E., Somaite,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,R., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,Y., Villalobos,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,  
 Weinstein,G., and Gibbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 184762)  
 Worley,K.C.  
 Direct Submission  
 Submitted (15-FEB-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 184762)  
 Worley,K.C.  
 Direct Submission  
 Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Aug 4, 2002 this sequence version replaced gi:22094227.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Drafting Center Code: BCM  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: HAFc  
 Center clone name: RP11-183H16  
 ----- Summary Statistics  
 Sequencing vector: plasmid;  
 Sequencing vector: M13;  
 Chemistry: dye-primer Bodypy: 2% of reads  
 Chemistry: dye-terminator Big Dye: 98% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 183845 bases at least Q40  
 Consensus quality: 184849 bases at least Q40  
 Consensus quality: 185447 bases at least Q20  
 Estimated insert size: 159026; sum-of-contigs estimation  
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 9 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1  
 \* 2041 contig of 2040 bp in length  
 \* 2141 gap of unknown length  
 \* 4170 contig of 2030 bp in length  
 \* 4270 gap of unknown length  
 \* 4171 gap of unknown length  
 \* 4271 contig of 4024 bp in length  
 \* 8295 gap of unknown length



8395 14196: contig of 5802 bp in length  
\* 14197 14296: gap of unknown length  
\* 14297 33076: contig of 18780 bp in length  
\* 33077 33176: gap of unknown length  
\* 33177 52295: contig of 1919 bp in length  
\* 52296 52395: gap of unknown length  
\* 52396 84185: contig of 31790 bp in length  
\* 84186 84285: gap of unknown length  
\* 84286 120023: contig of 35738 bp in length  
\* 120024 120123: gap of unknown length  
\* 120124 184762: contig of 64639 bp in length.  
\*  
FEATURES  
source  
1..184762  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-183H16"  
BASE COUNT 47425 a 44751 c 45101 g 46673 t 812 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 6,46e-36 Length: 184762  
Score: 472.00 Matches: 98  
Percent Similarity: 65.13% Conservative: 1  
Best Local Similarity: 64.47% Mismatches: 1  
Query Match: 49.22% Indels: 52  
Gaps: 2  
DB: 1  
US-10-082-502-19 (1-182) x AC023500 (1-184762)  
OY 69 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluValaCysAspArg 88  
|||||  
Db 43486 GTGCTTATGCGCCGTACAGAGCCACCTCACAGAGCTGCGAGAGATATGACCGG 43545  
|||||  
OY 89 MetTyrGluTyrGluGlnIleAspProSerThrHisArgLysAsnTyrValaArgVal 108  
|||||  
Db 43546 ATGAAGAGATGAGGGAACAAGTTATCTTCACACCATCGAAGACTACGACTGTA 43605  
|||||  
OY 109 ValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128  
|||||  
Db 43606 GTGGCCCGGAATGAGAACTCAGTGAACCTGACCTACAAAGCATCGAATCGACTCAGAT 43665  
|||||  
OY 129 IleSerGlyThrIleuLysPheAla----- 136  
|||||  
Db 43666 ATTACGGCACCTCTCAAGTTGCGGTGAGCTATGGAATCGGTACGTGCTTGGACAT 43725  
|||||  
OY 136 ----- 136  
Db 43726 TAAGGCTTTGGAGAACCAATTAGAAAGTTAGGCTGATATCTTGCCTCCTTTGG 43785  
|||||  
OY 136 ----- 136  
Db 43786 AAGTGAAGCAAGAGCCCTTATTCCTCTGTGCACCCCAATTCCTGGGTTGGCAG 43845  
|||||  
OY 137 CysGluSerIleValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAla 156  
|||||  
Db 43846 TGTGAGACATTTGTGAGAGATACAGAGATGACATCAATTCCTTTTCCGAGAGGCT 43905  
|||||  
OY 157 AspAsnValLysAspLysLeuGluSerLysArgThr 168  
|||||  
Db 43906 GACATATGTTAAAGACAACTTGCAGTAAAGGACAA 43941  
|||||  
RESULT 11  
AC090489 207424 bp DNA linear ROD 26-JAN-2002  
LOCUS AC090489  
DEFINITION Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence.  
AC090489  
AC090489  
AC090489.8 GI:18376849  
KEYWORDS HTG.  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
1 (bases 1 to 207424)  
McCombie,W.R., de la Bastide,M., Spiegel,L., Preston,R.,  
Kirchoff,K., Kuit,R., Nascimento,L., Zuberov,T., Balija,V.,  
Bell,M., Baker,J., Santos,L., Miller,B., Katzenberger,F.,  
Muller,S., King,L., Yang,C., Palmer,L., O'Shaughnessy,A. and  
Dehila,N.  
TITLE  
Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence  
JOURNAL  
REFERENCE  
2 (bases 1 to 207424)  
AUTHORS  
TITLE  
JOURNAL  
McCombie,W.R.  
Direct Submission  
Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
3 (bases 1 to 207424)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
McCombie,W.R.  
Direct Submission  
Submitted (26-JAN-2002) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
On Jan 26, 2002 this sequence version replaced gi:18201765.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest.  
FEATURES  
source  
1..207424  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-104010"  
/clone\_lib="RPCT-23"  
48132..48165  
misc-feature  
/note="We believe the assembly to be correct. The sequence  
is a mononucleotide (T) repeat in which the exact number  
of Ts is unknown. Other subclones in the region show one  
fewer T than that represented in the assembly."  
BASE COUNT 52613 a 48671 c 49663 g 56477 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.92e-34 Length: 207424  
Score: 451.50 Matches: 100  
Percent Similarity: 49.26% Conservative: 0  
Best Local Similarity: 49.26% Mismatches: 0  
Query Match: 47.08% Indels: 103  
Gaps: 1  
DB: 1  
US-10-082-502-19 (1-182) x AC090489 (1-207424)  
OY 69 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluValaCysAspArg 88  
|||||  
Db 75789 GTACCTTATGCGCCGTACAGAGCCACCTCACAGAGTGTGAGAGCTGTGACCGA 75848  
|||||  
OY 89 MetTyrGluTyrGluGlnIleAspProSerThrHisArgLysAsnTyrValaArgVal 108  
|||||  
Db 75849 ATGAAGAGATGAGGGAACAAGTTATCTTCACACCATCGAAGACTACGACTGTA 75908  
|||||  
OY 109 ValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 128  
|||||  
Db 75909 GTGAGCCGGAATGAGAGATCAAGTGAACCTTACACCCACCCGAAATGACTACAT 75968  
|||||  
OY 129 IleSerGlyThrIleuLysPheAla----- 136  
|||||  
Db 75969 ATCAGCGCACCTCCAAAGTTGCGGTGAGTTACCGCTGCGGAGAGTACTGTTCTTG 76028  
|||||  
OY 136 ----- 136  
Db 76029 GGAGTTATAGGATTCCTGCGCACAGATGGAGATGATGCGAATGTTGCTTAC 76088



OY 136 ----- 136  
 Db 76089 TTGTAACTTAGACACTCCAGGGGCGGTGGAGATCAGCTTAAGTTTGAGCCCTGGGC 76148  
 OY 136 ----- 136  
 Db 76149 CTAGTTCATAGTAGTTCGAGGCCAGCTGGGCAATATAGTAAGCCCTGTCTATTAATAAT 76208  
 OY 136 ----- 136  
 Db 76209 CCCCACCCACCCCAAAACAGTAGAGAGCTGAGCCAGCTGTCTTGGCCAGCTTGGCA 76268  
 OY 137 ----- 145  
 Db 76269 TCTGGCTCAATCTCTCTCTTCTTGGTCCGAGCTGAGCATCTGTGGAAGAATACGAG 76328  
 OY 146 AAGGJLleuIleGluPhePheSerArgJuaAAspAsnValIyAspIySer 165  
 Db 76329 GATGAGCTTATCGAATCTTCTCCAGAGAGGCTGACAACTTAAGACAAACTTGCAGT 76388  
 OY 166 LysArgThr 168  
 Db 76389 AAGCGGACA 76397

RESULT 12  
 AC109891 101666 bp DNA linear HTG 13-JUL-2002  
 LOCUS Rattus norvegicus clone CH230-319E6, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 47 unordered pieces.  
 AC109891  
 VERSION AC109891.3 GI:21738196  
 KEYWORDS HTG: HTGS\_PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 101666)

REFERENCE  
 AUTHORS  
 1 (bases 1 to 101666)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbarella,J., Benton,J., Blumerg,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowler,S., Brieleva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Barnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frenzt,P.,  
 Gabisl,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrill,J.H., Guevara,N., Guneratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Hawlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homel,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Korvah,J., Kovar,C.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kralovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,J., Liu,J.,  
 Lozodo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Mohabbat,K., Morgan,M., Morris,S.,  
 Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogub,M., Nguyen,N.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Peter,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shoshitani,N., Sisson,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Taboz,P., Tamerisa,A., Tamerisa,K., Tang,H.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Umanil,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleciyk,R., Woodson,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,Y., Zorrilla,S., Nelson,D.,  
 Weinstein,G., and Gibbs,R.  
 Direct Submission  
 2 (bases 1 to 101666)  
 Worley,K.C.  
 Direct Submission  
 Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 101666)  
 Worley,K.C.  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 12, 2002 this sequence version replaced g1:18860222.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GORY  
 Center clone name: CH230-319E6  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap version 0.990329  
 Consensus quality: 60228 bases at least Q40  
 Consensus quality: 63959 bases at least Q30  
 Consensus quality: 66730 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/GenBank\_drift\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 47 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1  
 \* 1423: contig of 1423 bp in length  
 \* 1424: gap of unknown length  
 \* 1524: contig of 1449 bp in length  
 \* 2972: gap of unknown length  
 \* 3072: gap of unknown length  
 \* 3073: gap of unknown length  
 \* 4717: contig of 1645 bp in length  
 \* 4718: gap of unknown length  
 \* 4818: gap of unknown length  
 \* 6632: contig of 1815 bp in length  
 \* 6733: gap of unknown length  
 \* 8040: contig of 1308 bp in length  
 \* 8140: gap of unknown length  
 \* 8141: gap of unknown length  
 \* 10044: contig of 1903 bp in length  
 \* 10144: gap of unknown length  
 \* 10144: gap of unknown length  
 \* 11238: gap of unknown length  
 \* 11238: gap of unknown length  
 \* 11238: gap of unknown length  
 \* 12362: contig of 1025 bp in length  
 \* 12363: gap of unknown length  
 \* 12463: gap of unknown length  
 \* 14101: contig of 1638 bp in length  
 \* 14201: gap of unknown length  
 \* 15840: contig of 1640 bp in length  
 \* 15940: gap of unknown length  
 \* 15941: gap of unknown length  
 \* 17054: contig of 1113 bp in length  
 \* 17153: gap of unknown length  
 \* 17154: gap of unknown length  
 \* 18836: contig of 1683 bp in length  
 \* 18936: gap of unknown length  
 \* 20137: contig of 1201 bp in length  
 \* 20237: gap of unknown length

20238 21736: contig of 1499 bp in length  
21737 21836: gap of unknown length  
21837 23619: contig of 1783 bp in length  
23620 23719: gap of unknown length  
23720 25907: contig of 2188 bp in length  
25908 26007: gap of unknown length  
26008 27406: contig of 1399 bp in length  
27407 27506: gap of unknown length  
27507 28837: contig of 1331 bp in length  
28838 28938 30229: contig of 1292 bp in length  
30230 30329: gap of unknown length  
30330 32012: contig of 1683 bp in length  
32013 32112: gap of unknown length  
32113 33198: contig of 1086 bp in length  
33199 33298: gap of unknown length  
33299 35519: contig of 2221 bp in length  
35520 35620 37677: contig of 2058 bp in length  
37678 37777: gap of unknown length  
37778 38936: contig of 1159 bp in length  
38937 39036: gap of unknown length  
39037 40273: contig of 1237 bp in length  
40274 40373: gap of unknown length  
40374 41906: contig of 1533 bp in length  
41907 42006: gap of unknown length  
42007 44506: contig of 2500 bp in length  
44507 44606: gap of unknown length  
44607 46803: contig of 2197 bp in length  
46804 46903: gap of unknown length  
46904 48226: contig of 2323 bp in length  
48227 48326: gap of unknown length  
48327 51433: contig of 2107 bp in length  
51434 51533: gap of unknown length  
51534 53763: contig of 2230 bp in length  
53764 53863: gap of unknown length  
53864 56613: contig of 2750 bp in length  
56614 56713: gap of unknown length  
56714 58773: contig of 2060 bp in length  
58774 58873: gap of unknown length  
58874 61527: contig of 2654 bp in length  
61528 61627: gap of unknown length  
61629 63192: contig of 1565 bp in length  
63193 63292: gap of unknown length  
63293 65362: contig of 2070 bp in length  
65363 65462: gap of unknown length  
65463 67438: contig of 1976 bp in length  
67439 67538: gap of unknown length  
67539 70505: contig of 2967 bp in length  
70506 70605: gap of unknown length  
70606 74524: contig of 3919 bp in length  
74525 74624: gap of unknown length  
74625 76737: contig of 2113 bp in length  
76738 76837: gap of unknown length  
76839 80133: contig of 3296 bp in length  
80134 80233: gap of unknown length  
80234 83122: contig of 2889 bp in length  
83123 83222: gap of unknown length  
83223 86731: contig of 3509 bp in length  
86732 86831: gap of unknown length  
86832 89929: contig of 3098 bp in length  
89930 90029: gap of unknown length  
90030 93107: contig of 3078 bp in length  
93108 93207: gap of unknown length  
93208 97419: contig of 4212 bp in length  
97420 97519: gap of unknown length  
97520 101666: contig of 4147 bp in length.

FEATURES  
source

1..101666  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-319E6"  
24650 a 22934 c 23569 g 23059 t 7454 others

## ORIGIN

## Alignment Scores:

Pred. No.: 7,06e-32 Length: 101666  
Score: 429.00 Matches: 97  
Percent Similarity: 49.25% Conservative: 1  
Best Local Similarity: 48.74% Mismatches: 2  
Query Match: 44.73% Indels: 99  
DB: 2 Gaps: 1

US-10-082-502-19 (1-182) x AC109891 (1-101666)

QY 69 VALPROTYRALAARGSERGLUALAHISLEUTHRGULEULEUGLUVALCYSAAPARG 88  
DB 97572 GTTCCCTTATGCCCTCCAGAGCCCACTCACGGAGTTGCTTGAAGAGATATGTACCGA 97631  
QY 89 MELLTSGUUTRGVLGVLGVLINLEASPROSETRHISATGLYSASTYRVALARGVAL 108  
DB 97632 ATGAAAGAGATAGGGGAACAGATTGACCTTCTACCCACCGCAACACTACCTCCGTCTT 97691  
QY 109 VALSERARGASNGLYGLUSERSERGLUENUSAPLEUGINGLYILEARGILEASPERASP 128  
DB 97692 GTCCGCCGGAACGAGAGATCCAGTACACTACAGGCGTATCCGAATGATTCAGAT 97751  
QY 129 ILESERGLYTHYLEULYPHE----- 135  
DB 97752 ATCACTGGCACCCCTCAAGATTGCGGTGAGTATGCGTGCAGTACAGTGTCTTCTG 97811  
QY 135 ----- 135  
DB 97812 GGAGTTCAATAGGAGATCTCGGCACACAGATTGGAGTCCGGCGGATATGTGTTCATG 97871  
QY 135 ----- 135  
DB 97872 CTTGTAACCGGCGAGTGAGAGAGCAGTGTGAAGATTGGCACTAGTTCATAGT 97931  
QY 135 ----- 135  
DB 97932 TCTAGCGCAGCTGAGCATATATCAATAAACCTGTCTATAAATCCACATGCCACCA 97991  
QY 135 ----- 135  
DB 97992 AAACGAGGAGAGACTGAGCCGCTCTCTGCCCCGCTGAGATCCTGGTCACTTCT 98051  
QY 136 -----Ala-CysGluSerIleValGluIuTyrGluAspGluIleG1 150  
DB 98052 CTGTCTTCTGTGTCGCGCAGTGTGAGAGCATTTGTGAGAGAGTGAACCTTATGCA 98111  
QY 150 uPhePheSerArgGluAlaAspAsnValLysAspLysLeuCysserLysArgThr 168  
DB 98112 ATCTTTTCAAGAGAGGCTGACCAACGTTAAAGACAACCTTGCACTAAGCGGACA 98166

RESULT 13  
AF186113 649 bp mRNA linear PRI 13-JAN-2000  
LOCUS Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds.  
DEFINITION AF186113  
ACCESSION AF186113  
VERSION AF186113.1 GI:6014631  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 649)  
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.  
TITLE Homo sapiens putative secreted protein  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 649)  
AUTHORS Sheppard,P., Jelinek,L., Whitmore,T., Blumberg,H., Lehner,J. and O'Hara,P.  
TITLE Direct Submission



SOURCE

Norway rat.

ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 69753)  
Munry,D.M., Adams,C., Adio-Odule,B., Ali-osman,F.R., Allen,C.,  
Albrooks,S.L., Amaratunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbala,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P.,  
Bulay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinn,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorell,J., H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsom,B., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
Lozano,R., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Mashevari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Monabhat,K., Morgan,M., Morris,S.,  
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okunolu,G.,  
Otaguine,N., Oviedo,R., Pace,A., Payton,B., Peary,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherrer,S., Scott,G., Shen,H., Shooshtari,N., Sison,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,U., Zorrilla,S., Nelson,D.,  
Welnstock,G. and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL

Unpublished  
2 (bases 1 to 69753)  
Direct Submission  
Submitted (24-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 69753)  
Worley,K.C.

REFERENCE  
AUTHORS  
JOURNAL

Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 11, 2002 this sequence version replaced gi:17974643.

COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: GJIX  
Center clone name: CH230-208H17  
Sequencing Statistics  
Chemistry: Dye-terminator Big Dye 1000 of reads  
Assembly program: Phrap, version 0.990329

Consensus quality: 15506 bases at least Q40  
Consensus quality: 17125 bases at least Q30  
Consensus quality: 18271 bases at least Q20

NOTE: Estimated insert size may differ from sequence length.  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drift\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)).  
NOTE: This is a working draft sequence. It currently  
consists of 42 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1  
1366: contig of 1366 bp in length  
1466: gap of unknown length  
2632: contig of 1166 bp in length  
2732: gap of unknown length  
3919: contig of 1187 bp in length  
4019: gap of unknown length  
5169: contig of 1150 bp in length  
5269: gap of unknown length  
6813: contig of 1544 bp in length  
6913: gap of unknown length  
8119: contig of 1206 bp in length  
8219: gap of unknown length  
9384: contig of 1165 bp in length  
9484: gap of unknown length  
10996: contig of 1512 bp in length  
11096: gap of unknown length  
12554: contig of 1458 bp in length  
12555: gap of unknown length  
14376: contig of 1722 bp in length  
14476: gap of unknown length  
15790: contig of 1314 bp in length  
15890: gap of unknown length  
17010: contig of 1120 bp in length  
17110: gap of unknown length  
18308: contig of 1198 bp in length  
18408: gap of unknown length  
19786: contig of 1378 bp in length  
19886: gap of unknown length  
21145: contig of 1259 bp in length  
21245: gap of unknown length  
22539: contig of 1294 bp in length  
22639: gap of unknown length  
23861: contig of 1222 bp in length  
23961: gap of unknown length  
25395: contig of 1434 bp in length  
25495: gap of unknown length  
26515: contig of 1020 bp in length  
26615: gap of unknown length  
28329: contig of 1714 bp in length  
28429: gap of unknown length  
30046: contig of 1617 bp in length  
30146: gap of unknown length  
31613: contig of 1467 bp in length  
31713: gap of unknown length  
33120: contig of 1407 bp in length  
33220: gap of unknown length  
34615: contig of 1395 bp in length  
34715: gap of unknown length  
35867: contig of 1152 bp in length  
35967: gap of unknown length  
37671: contig of 1704 bp in length  
37771: gap of unknown length  
39476: contig of 1705 bp in length  
39576: gap of unknown length  
41221: contig of 1645 bp in length  
41321: gap of unknown length  
43130: contig of 1809 bp in length  
43230: gap of unknown length  
44734: contig of 1504 bp in length





GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1535.7 Seconds

(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-20

Perfect score: 848  
Sequence: 1 RRSDDLHCACRALVDELEW.....LCKRPTDLDHALHSHDEL 162

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+p2n.model -DEV=xlp  
-O=/gen2.1/USPTO\_spool/US10082502/runat.24032003.135059.6547/app.query.fasta.1.1308  
-DB=GenEmbl -QFWT=fastap -SUFFIX=rg -MINMATCH=0.1 -IOOFC=0 -IOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=pc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502 -CGEN=1.1.4758 -runat.24032003.135059.6547 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_XMAP -LARGEQUER -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_em:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pin:\*  
35: em\_hcg\_fod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hgo\_hum:\*  
40: em\_hgo\_mus:\*  
41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	848	100.0	814	6	AX440456
2	848	100.0	814	9	AB015631
3	848	100.0	1080	9	AY032624
4	848	100.0	1210	6	AX464040
5	830	97.9	798	10	BC008261
6	830	97.9	1017	10	AF186115
7	830	97.9	158198	2	AC012013
8	488.5	57.6	155023	2	AC025574
9	481	56.7	182350	2	AC073896
10	481	56.7	184762	2	AC023500
11	446.5	52.7	207424	10	AC090489
12	436	51.4	101666	2	AC109891
13	257	30.3	649	9	AF186113
14	257	29.0	2378	9	BC001027
15	246	26.5	69753	2	AC103156
16	224.5	26.5	996	3	AY094698
17	202	23.8	996	3	U13070
18	190.5	22.5	40824	5	AB055671
19	155.5	18.3	622	5	AB055671
20	154	18.2	564	6	AY088187
21	154	18.2	785	8	AY088187
22	141.5	16.7	175963	3	AC009379
23	141.5	16.7	279530	3	AE003518
24	141	16.6	55359	2	AC012854
25	141	16.6	172804	3	AC007414
26	141	16.6	275390	3	AE003831
27	138.5	16.3	1286	3	AY060987
28	131	15.4	161566	9	AC008060
29	129	15.2	744	6	AR142811
30	129	15.2	746	6	AX136616
31	129	15.2	747	6	BD006701
32	129	15.2	1413	9	BC019903
33	129	15.2	1419	6	AX136423
34	129	15.2	1470	6	AX376412
35	129	15.2	1470	6	AX454716
36	129	15.2	1470	6	AX491194
37	129	15.2	1512	6	AR142810
38	129	15.2	1547	6	BC032339
39	124	14.6	76187	2	AC015219
40	116	13.7	168418	9	AC009403
41	115	13.6	181896	2	AC114126
42	115	13.6	201649	2	AC120813
43	114.5	13.5	201649	2	AC120813
44	113	13.3	1260	6	AX380553
45	111	13.1	210631	2	AC129184

RESULT 1

## ALIGNMENTS



AX440456  
 LOCUS AX440456 814 bp DNA linear PAT 28-JUN-2002  
 DEFINITION Sequence 309 from Patent WO0190154.  
 ACCESSION AX440456  
 VERSION AX440456.1 GI:21665266  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS 1  
 Xu, J., Mitcham, J. L., Harlocker, S. L., Dillon, D. C., Secrist, H.,  
 Lodes, M. J., Algate, P. A., Fling, S. P., Mannion, J., Benson, D. R. and  
 Carter, D.  
 TITLE Compositions and methods for the therapy and diagnosis of ovarian  
 cancer  
 JOURNAL Patent: WO 0190154-A 309 29-NOV-2001;  
 CORIXA CORPORATION (US)  
 FEATURES  
 SOURCE 1. 814  
 Location/Qualifiers  
 BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.92e-77 Length: 814  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-082-502-20 (1-162) x AX440456 (1-814)

OY 1 ArgATgSerGlnAspLeuHnIScYsGlyAlaCysATgAlaLeuValAspGluLeuGluTrp 20  
 |||||||  
 DB 205 CGAGAGCAGCAGATCTCCACTGTGAGACATGACAGCGCTGTGGATGAACTAGAAATGG 264  
 OY 21 GluIleAlaGlnValAspProLySylThrIleGlnMetGlySerPheArgIleAsnPro 40  
 |||||||  
 DB 265 GAAATGGCCAGCTGAGCCCAAGAACATTCAGATGGGATCTTCCGATCAATCCA 324  
 OY 41 AspGlySerGlnSerValAlaGluValProTyraIlaArgSerGluAlaHnIleuThrGlu 60  
 |||||||  
 DB 325 CATGACAGCAGCTCAGTGTGGAGGTGCTTATGCCCCCTCAGAGCCCACTCAGAG 384  
 OY 61 LeuLeuGluGluIleCysAspArgMetLySyluTyrgIyGluGlnIleAspProSerThr 80  
 |||||||  
 DB 385 CTGCTGGAGGAGATATGTACCCGATGAAGAGTATGGGGAACATTTGATCTCCACC 444  
 OY 81 HIsArgLySAsnTyrValAlaGlyValAlaIyArgAsnGlyGluSerSerGluLeuAspLeu 100  
 |||||||  
 DB 445 CATCCCAAGAACTACGTACGTACTGTGCGCGGAGATGAGAAATCCAGTAACCTGACCTTA 504  
 OY 101 GlnGlyIleArgTyIleAspSerAspIleSerGlyThrLeuLySPhaIleCysGluSerIle 120  
 |||||||  
 DB 505 CAAAGCAGATCCGATGACTCAATATTAAGCGCACCTTCAGTTGGCTGTGAGAGCATT 564  
 OY 121 ValGluGluTyrgIuAspGluLeuIleGluPhePheSerATgGluAlaAspAsnValLyS 140  
 |||||||  
 DB 565 GTGGAGGAATAGCAGAGTAAGTAATCTTTTCCGAGAGCGTACACATTTTAA 624  
 OY 141 AspLySLeuCySserLyArgThrAspLeuCySAspHnIleAlaLeuHnIleSerHnIAsp 160  
 |||||||  
 DB 625 GACAAACTTTCAGTAAGCAAGACAGATCTTTGTGCACATGCCCTGCACATATGCGATGAT 684  
 OY 161 GluLeu 162  
 |||||||  
 DB 685 GAGCTA 690

RESULT 2  
 AB015631

AB015631  
 LOCUS AB015631 814 bp mRNA linear PRI 26-JUL-2001  
 DEFINITION Homo sapiens mRNA for type II membrane protein, complete cds,  
 clone:HP10390.  
 ACCESSION AB015631  
 VERSION AB015631.1 GI:4586839  
 KEYWORDS type II membrane protein.  
 SOURCE Homo sapiens gastric adenocarcinoma cDNA to mRNA,  
 clone\_11b:pkal-meta-1 clone:HP10390.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS 1  
 Yokoyama-Kobayashi, M., Yamaguchi, T., Sekine, S. and Kato, S.  
 TITLE Selection of cDNAs encoding putative type II membrane proteins on  
 the cell surface from a human full-length cDNA bank  
 JOURNAL Gene 228 (1-2), 161-167 (1999)  
 MEDLINE 99173880  
 REFERENCE 2 (bases 1 to 814)  
 AUTHORS Kato, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JUN-1998) Selshi Kato, Research Institute of National  
 Rehabilitation Center for the Disabled, Department of  
 Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama  
 359-8555, Japan (E-mail: selshi@rehab.go.jp,  
 Tel:042-995-3100(ex.2568), Fax:042-995-3132)

FEATURES  
 SOURCE 1. 814  
 Location/Qualifiers  
 CDS  
 1. 814  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="HP10390"  
 /tissue\_type="gastric adenocarcinoma"  
 /clone\_11b="pkal-meta-1"  
 145..693  
 /codon\_start=1  
 /product="type II membrane protein"  
 /protein\_id="PAA76498.1"  
 /db\_xref="GI:4586840"  
 /translation="MKSGKVALIGLALGILGTRARRSODLHCACRALVDELEETIQ  
 VDKRTIOMSEFRINPDSOSVVEVPARSANHTLEELICDMMKRYGQIDPSTIR  
 KNVYRVGRNGESSELDIGIRIDSDISGLTKFACESIVEREDELIEFESREADNVK  
 DKLSKRTDLCDHALHSHDEL"

BASE COUNT 210 a 194 c 241 g 169 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.92e-77 Length: 814  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x AB015631 (1-814)

OY 1 ArgATgSerGlnAspLeuHnIScYsGlyAlaCysATgAlaLeuValAspGluLeuGluTrp 20  
 |||||||  
 DB 205 CGAGAGCAGCAGATCTCCACTGTGAGACATGACAGCGCTGTGGATGAACTAGAAATGG 264  
 OY 21 GluIleAlaGlnValAspProLySylThrIleGlnMetGlySerPheArgIleAsnPro 40  
 |||||||  
 DB 265 GAAATGGCCAGCTGAGCCCAAGAACATTCAGATGGGATCTTCCGATCAATCCA 324  
 OY 41 AspGlySerGlnSerValAlaGluValProTyraIlaArgSerGluAlaHnIleuThrGlu 60  
 |||||||  
 DB 325 CATGACAGCAGCTCAGTGTGGAGGTGCTTATGCCCCCTCAGAGCCCACTCAGAG 384  
 OY 61 LeuLeuGluGluIleCysAspArgMetLySyluTyrgIyGluGlnIleAspProSerThr 80  
 |||||||  
 DB 385 CTGCTGGAGGAGATATGTACCCGATGAAGAGTATGGGGAACATTTGATCTCCACC 444  
 OY 81 HIsArgLySAsnTyrValAlaGlyValAlaIyArgAsnGlyGluSerSerGluLeuAspLeu 100  
 |||||||  
 DB 445 CATCCCAAGAACTACGTACGTACTGTGCGCGGAGATGAGAAATCCAGTAACCTGACCTTA 504

QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 505 CAAGGATCCGAAATCGACATGATATTACGGACACCTCAAGTTCCGCTGAGACATTT 564  
QY 121 ValGluGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 565 GTGGAGAGATACGAGATGACATCATTTGATTTTCCGAGAGAGCTGACATGTTTAA 624  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 625 GACAACTTTGACATGAGGAGACAGATCTTTGTGACACATCCCTGACATATGCATGAT 684  
QY 161 GluLeu 162  
DB 685 GAGCTA 690

RESULT 3  
AY032624 1080 bp mRNA linear PRI 18-APR-2002  
LOCUS AY032624  
DEFINITION Homo sapiens saposin-like protein mRNA, complete cds.  
ACCESSION AY032624  
VERSION AY032624.1 GI:20196198  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1080)  
AUTHORS Bornhauser,B.C., Olsson,P.-A. and Lindholm,D.  
TITLE NSAP is a novel saposin-like protein that interacts with MIR and stimulates neurite outgrowth  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1080)  
AUTHORS Olsson,P.-A. and Lindholm,D.  
TITLE Direct Submission  
JOURNAL Submitted (13-APR-2001) Neuroscience, Uppsala University,  
Husaratan 3, Uppsala 75123, Sweden  
FEATURES  
SOURCE location/Qualifiers  
1. 1080  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q13"  
419..967  
/note="NSAP; encodes a type 2 membrane signal and a putative ER retention motif"  
/codon\_start=1  
/product="saposin-like protein"  
/protein\_id="AAK38148.1"  
/db\_xref="GI:20196198"  
/translation="MKGMWLALLGLGTAHARRSODLHGACRAIVDELEETIAQ  
VDPKTIQMGSEFRINPDGOSVVEVYARSEAHLELEETCDPKKEVGEQIDSTHR  
KNYVRVGRNGSESLDLOGIRIDSDISGTLKFACESTVEYEDLEIFESKEDNVK  
DKTSKRTDLCDAHLHSHDEL"

BASE COUNT 238 a 269 c 343 g 230 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5,53e-77 Length: 1080  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-20 (1-162) x AY032624 (1-1080)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20  
DB 479 CGGAGAGCCAGATCTCCACCTGTGAGAGATCGAGGCTCTGTGTGAGTGAATGAG 538  
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40

DB 539 GAAATGGCCAGGTGGACCCCAAGAACATTCAGATGGATCTTCCGATCAATCCA 598  
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 599 GATGGACACCACTGCTGAGTGTGGAGGTGCTTATCCCGCTCAGAGGCCACCTCAGAG 658  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 659 CTGCTGGAGAGATATGACCGGATGAGAGATATGGGACACAGATTCCTTCACCC 718  
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100  
DB 719 CATGCCAAGACTACGTACGTGTGTGGCCGAGATGAGAAATCCAGTAACTGAGACCTA 778  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 779 CAAGGATCCGAAATCGACATGATATTACGGACACCTCAAGTTCCGCTGAGACATTT 838  
QY 121 ValGluGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 839 GTGGAGAGATACGAGATGAGTGAATCATTTCTTTCCGAGAGGCTGACATGTTTAA 898  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
DB 899 GACAACTTTGACATGAGGAGACAGATCTTTGTGACACATCCCTGACATATGCATGAT 958  
QY 161 GluLeu 162  
DB 959 GAGCTA 964

RESULT 4  
AX464040 1210 bp DNA linear PAT 16-JUL-2002  
LOCUS AX464040  
DEFINITION Sequence 173 from Patent WO0140466.  
ACCESSION AX464040  
VERSION AX464040.1 GI:21899037  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Baker,K.P., Beresini,M., DeForge,L., Desnoyers,L., Flivetroff,E.,  
Gao,M.O., Gerlitsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,  
Sherwood,S., Smith,Y., Stewart,T.A., Tumas,D., Wetlanbe,C.K.,  
Wood,W.L. and Zhang,Z.  
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding  
JOURNAL same  
JOURNAL Patent: WO 0140466-A 173 07-JUN-2001;  
Genentech Inc. (US)  
FEATURES  
SOURCE location/Qualifiers  
1..1210  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 261 a 316 c 379 g 254 t  
ORIGIN

Alignment Scores:  
Pred. No.: 6,35e-77 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-20 (1-162) x AX464040 (1-1210)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20  
DB 603 CGGAGAGCCAGATCTCCACCTGTGAGAGATCGAGGCTCTGTGTGAGTGAATGAG 662  
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40

|||||  
Db 663 GNAATGGCCAGGTGAGCCCAAGAGACCATTCAGATGGATCTTCCGGATCAATCCA 722  
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
Db 723 GATGACACCAAGTCAAGTGTGAGGTGCTTATGCCCGCTCGAGAGCCCACTCACAGAG 782  
Oy 61 LeuLeuGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80  
Db 783 CTGCTGGAGGAGATATGATGACCGGATGGAAGGATGAGGAAACAGATTCCTCCAC 842  
Oy 81 HisArgLysAsnTyrValArgValValGluValArgAsnGlyLysSerGlnLeuAspLeu 100  
Db 843 CATCCCAAGAACTACGTCATGCTGAGTGGCCGGAATGGAATCCAGTGAATGGACCTA 902  
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
Db 903 CAGGCAATCCGAATGAGTCAAGATATTAGCGGCACCCCAAGTTGGCTGTAGAGCAT 962  
Oy 121 ValGluGluTyrGluAspGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140  
Db 963 GTGCGAGATACGAGATGAACTCATTCATTCCTTCCCGAGAGCTGACAAATCTTAAA 1022  
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisLeuSerHisAsp 160  
Db 1023 GCAAAACTTTCAGTACGACGACAGATCTTGTGACATGCCCCGACATATGCGATGAT 1082  
Oy 161 GluLeu 162  
Db 1083 GAGCTA 1088

RESULT 5  
BC008261 798 bp mRNA linear ROD 07-AUG-2002  
LOCUS  
DEFINITION Mus musculus, transmembrane protein 4, clone MGC:6853  
ACCESSION BC008261  
VERSION BC008261.1 GI:14198400  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 798)  
Strausberg, R.  
Direct Submission  
Submitted (22-MAY-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

REMARK  
COMMENT

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 5 Row: B Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 9903606.  
Location/Qualifiers  
1..798  
/organism="Mus musculus"

/db\_xref="taxon:10090"  
/map="129, C57BL/6J, FVB/N"  
/clone="MGC:6853 IMAGE:2650612"  
/tissue\_type="Mammary tumor. Brca1-/fl; MMTV-Cre model. 10  
months old, gross tissue."  
/clone\_lib="NCI\_CGAP\_Mam3"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
115..663  
/codon\_start=1  
/product="transmembrane protein 4"  
/protein\_id="AAH08261.1"  
/db\_xref="GI:14198401"  
/db\_xref="LocusID:56530"  
/translation="MGWMLALLGLVLTAMARRSODLHCGRALVDELEIAR  
VDPKTYMGSEFRINPDSOSVPEPARSAHLLELEVCDBMKKEGEOIDPSTHR  
KNYRVVSRNSESSELDIGTRIDSDISGTLKPFACSEIVEYEDBELIEFSEADNWK  
DKLSKRTDLCDHALHSHDEL"  
BASE COUNT 227 a 182 c 228 g 161 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,56e-75 Length: 798  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 0  
Query Match: 97.88% Indels: 0  
DB: 10 Gaps: 0

US-10-082-502-20 (1-162) x BC008261 (1-798)

Oy 1 ArgArgSerGlnAspLeuHisCysGluValAcysArgAlaLeuValAspGluLeuGluTrp 20  
Db 175 CGAAGGAGCCCAAGATCTACACTGTGAGCTTGACAGGCTGTGAGTGAATTAAGATGG 234  
Oy 21 GluLeuAlaGlnValAspProLysThrIleGlnMetGlySerPheArgGlnAsnPro 40  
Db 235 GAAATTCGCCGCGGTGAGCCCAAGAGACCATTCAGATGGATCCCTCCGAATCAATCCA 294  
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
Db 295 GATGACACCAAGTCAAGTGTGAGGTACTTATGCCCGCTCGAGAGCCCACTCACAGAG 354  
Oy 61 LeuLeuGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80  
Db 355 TTGCTTGAAGAGGTGTGTGACCGAATGAAGGATGACGGGAAACAGATTCACCTTAC 414  
Oy 81 HisArgLysAsnTyrValArgValValGluValArgAsnGlyLysSerGlnLeuAspLeu 100  
Db 415 CACCGCAAGAACTACGTCATGCTGAGCCGGAATGGAATCCAGTGAATGACTTA 474  
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
Db 475 CAGGCAATCCGAATGAGTCAAGATATTAGCGGCACCCCAAGTTGGCTGTAGAGCAT 534  
Oy 121 ValGluGluTyrGluAspGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140  
Db 535 GTGCGAGATACGAGATGAACTCATTCATTCCTTCCCGAGAGCTGACAAATCTTAAA 594  
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisLeuSerHisAsp 160  
Db 595 GACAAACTTTCAGTACGACGACAGATCTATGTGACCATGCCCTGACACAGATCTCAGAT 654  
Oy 161 GluLeu 162  
Db 655 GAGCTA 660

RESULT 6  
AF186115 1017 bp mRNA linear ROD 13-JAN-2000  
LOCUS AF186115  
DEFINITION Mus musculus putative secreted protein zslg9 (zslg9) mRNA, complete  
cde.  
ACCESSION AF186115

VERSION AF186115.1 GI:6014635  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Mus musculus putative secreted protein  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1017)  
 AUTHORS Sheppard, P., Jellinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
 O'Hara, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-SEP-1999) Biomolecular Informatics, Zymogenetics,  
 Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
 FEATURES  
 SOURCE  
 1..1017  
 /organism="Mus musculus"  
 /db\_xref="dbEST:AA049839"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:476061"  
 1..1017  
 /gene="Zs1g9"  
 343..891  
 /gene="Zs1g9"  
 /codon\_start=1  
 /product="putative secreted protein ZS1g9"  
 /protein\_id="AF01433.1"  
 /db\_xref="GI:6014636"  
 /translation="MKKGMWALLLVLTATARRSODLHCACALYDELEMETAR  
 VDPKTIOMGSEFRINDGSGVEVYAPYASRALTLELEVCNMEYEDQIDPSTHR  
 KNYVRVNRNGSESRDLGIRIDSDISGLTFACRSIVEEYEDLEIFESREADNVK  
 DKTSRTDCAHLRSHDEL  
 250 c 315 g 204 t  
 BASE COUNT  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,44e-75 Length: 1017  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 2  
 Query Match: 97.88% Indels: 0  
 Gaps: 0  
 US-10-082-502-20 (1-162) x AF186115 (1-1017)  
 QY 1 ARGARGSGINAPLEUHNHSCYGYALACYSARGALALEUVALASPGIUEUGIUTYP 20  
 |||||  
 DB 403 CGAAGGCGCAAGATCTACACTGTGAGCTTGCGAGGCGCTGCTGATGATTAGATGG 462  
 QY 21 GIIIEALAGINVALAPPROLYSLYSTRILLEGIMETGLYSEPHARGILEANPRO 40  
 |||||  
 DB 463 GAAATGGCCGGGTGACCCCAAGAGACATTTCAGATGGATCCCTCCGAATCAATCCA 522  
 QY 41 ASPEGLYSERGLISERVALVALGLUVALPROTYRALAARGSERGLUALAHISLEUTHRI 60  
 |||||  
 DB 523 GATGGGAGCCAGTCACTTGTGAGGACTTATGCCCGCTGAGAGCCCACTCACAGG 582  
 QY 61 LEULEGLIGIUIECYSAPARGMETLVSGLIUTRYGILGUGIINLEASPPOSETHR 80  
 |||||  
 DB 583 TTGCTTGAGAGAGGTGTGTACCGCAATGAGAGGTACGGGGAACAGATTGACCCCTTAC 642  
 QY 81 HISAAGLVASANTYVALARGVALVALGLYARGAANGILYGLUSERSERGLIUEANSPLEU 100  
 |||||  
 DB 643 CACCCCAAGAACTACGTACGCTGTGAGCGGAGATGAGATCCAGTAACACTTA 702  
 QY 101 GINGIYIIEARGIIEASERASPLISERGLYTHRIEULYSPHALACYSGLUSERILE 120  
 |||||  
 DB 703 CAGGCGATCCGATTCAGATTCAGAGCGGACCCCAAGTTGCTGTGAGAGCATT 762  
 QY 121 VALGIUGIUTRYGILUASPGIUEULIEGLIUPHESERARGIUALAASPANVALYS 140

DB 763 GTGAGAGAAATACGAGAGGAGACTTATCGAATTCCTTCACAGAGGCTCGACCACTTTAAA 822  
 QY 141 ASPLYSLCUCYSSELYTARGTTHRASPLEUCYASPHISALALEUHNHISLESERHISASP 160  
 DB 823 GACAAACCTTTCAGTAAAGCGAGCAATCTATGTACCATGCCCTCCGACAGATCTACAGAT 882  
 QY 161 GLIUEU 162  
 |||||  
 DB 883 GAGCTA 888  
 RESULT 7  
 AC012013  
 LOCUS  
 DEFINITION  
 Homo sapiens chromosome 12 clone RP11-764L14, WORKING DRAFT  
 SEQUENCE, 5 unordered pieces.  
 AC012013  
 AC012013.18 GI:14717292  
 HTG: HTGS\_PHASE1, HTGS\_DNAFT, HTGS\_ACTIVEFIN.  
 Homo sapiens  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 158198)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
 Alshrocks, S.L., Amaratunga, H.C., Are, J.R., Banks, P.R., Benton, J.,  
 Benton, J., Blum, K., Blankenburg, K., Bonin, D., Bouch, J.,  
 Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,  
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F.,  
 Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,  
 Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
 Coyle, M.D., Dethorne, S.R., David, R., Davila, M.L., Davis, C.,  
 Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
 Den, A.L., Ding, X., Dinn, H.H., Douthett, C., Edgar, D., Edwards, C.C.,  
 Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edger, D., Edwards, C.C.,  
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
 Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,  
 Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S.,  
 Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A.,  
 Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,  
 Hollins, B., Homs, F., Howard, S., Huber, J., Hui, S., Hume, J.,  
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
 Joudan, S., Karlsson, E., Kelly, S., Khan, U., Kling, L., Korah, J.,  
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
 Lewis, L., Li, J., Li, Z., Lichtenberg, O., Lieu, C., Liu, D., Liu, W.,  
 Louisedge, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,  
 Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
 Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nokenkwo, S.,  
 Ogih, M., Okunou, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B.,  
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.,  
 Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shooshbari, N.,  
 Slisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,  
 Stone, H., Sutton, A., Swalek, A., Taber, P., Tamerisa, A., Tamerisa, K.,  
 Tang, H., Tansey, J., Taylor, C., Taylor, P., Telford, B., Thomas, N.,  
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,  
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
 Wellington, S., Williams, G., Williamson, A., Wleczky, R., Woodson, S.,  
 Worley, K., Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 158198)  
 Worley, K.C.  
 TITLE  
 JOURNAL  
 COMMENT  
 Submitted (19-OCT-1999) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 On Jul 13, 2001 this sequence version replaced g1:14547727.  
 ----- Genome Center



TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 155023)  
AUTHORS Worley, K. C.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 155023)  
Worley, K. C.  
REFERENCE Direct Submission  
AUTHORS Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 3, 2002 this sequence version replaced g1:20335511.  
JOURNAL Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: HALV  
Center clone name: RP11-348M3  
Summary Statistics  
Sequencing vector: M13  
Chemistry: Dye-Primer Bodypy: 5% of reads  
Chemistry: Dye-Terminator Big Dye: 95% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 141522 bases at least Q40  
Consensus quality: 147097 bases at least Q40  
Consensus quality: 150442 bases at least Q20  
Estimated insert size: 150706; sum-of-coverage estimation  
Quality coverage: 4x in Q20 bases; sum-of-coverage estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a working draft sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 3033: contig of 3033 bp in length  
3034 3133: gap of unknown length  
3134 5971: contig of 2838 bp in length  
5972 6071: gap of unknown length  
6072 10270: contig of 4198 bp in length  
10270 10369: gap of unknown length  
10369 13009: contig of 2640 bp in length  
13009 13109: gap of unknown length  
13109 17573: contig of 4464 bp in length  
17573 17673: gap of unknown length  
17673 24650: contig of 6977 bp in length  
24650 24751: gap of unknown length  
24751 31778: contig of 7028 bp in length  
31778 31878: gap of unknown length  
31878 42263: contig of 10385 bp in length  
42263 42364: gap of unknown length  
42364 51104: contig of 8741 bp in length  
51104 51205: gap of unknown length  
51205 68306: contig of 17102 bp in length  
68306 68407: gap of unknown length  
68407 109481: contig of 41075 bp in length  
109481 109582: gap of unknown length  
109582 155023: contig of 45442 bp in length.

FEATURES  
SOURCE  
1. 155023  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-348M3"

BASE COUNT 40008 a 36207 c 35987 g 41693 t 1128 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 7 01e-38 Length: 155023  
Score: 488.50 Matches: 114  
Percent Similarity: 39.72% Conservative: 0  
Best Local Similarity: 39.72% Mismatches: 0  
Query Match: 57.61% Indels: 173  
Gaps: 2 2  
DB: 2 2

US-10-082-502-20 (1-162) x AC025574 (1-155023)

QY 49 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluLeuIleCysAspArg 68  
|||||  
DB 56120 GTGGCTTATGCCCGCCGTCAGAGGCCACCTCAGACAGCTCTGGAGAGATATGTGACCG 56061  
|||||  
QY 69 MetLysGluTyrGlyGluGluIleAspProSerThrHisArgLysAsnTyrValaVal 88  
|||||  
DB 56060 ATCAAGAGATATGGGAGACAGATTCCTCCACCCATCGCAAGAACTACGATCGTGA 56001  
|||||  
QY 89 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlyTyrLeuArgTyrLeuAspAsp 108  
|||||  
DB 56000 GTGGGCCGAATGGAGAAATCCAGTGAATGACCTACAAAGCATCCGAATCCAGTCAGAT 55941  
|||||  
QY 109 IleSerGlyThrLeuLysPheAla 116  
|||||  
DB 55940 ATTACGGGACCCCTCAAGTTTGGGGTGAGCTATGGGAATCGGAGCTGTTTGACAT 55881  
|||||  
QY 116 116  
DB 55880 TAAAGCGTTGTTGAGAACCAATTAGAAATGAGCTGATTCCTTCCTCTTTGG 55821  
|||||  
QY 116 116  
DB 55820 AAGTGAGCGAAGAGCCCTTATTCCTGCTGTCACCCATTTTCCTGGTTGGCAG 55761  
|||||  
QY 117 CysGluSerTyrValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAla 136  
|||||  
DB 55760 TTGAGAGAGATTTGTGAGGAATTCAGAGATGAACTATTAATCTTTTCCGAGAGGCT 55701  
|||||  
QY 137 AspAsnValLysAspLysLeuCysSerLysArgThr 148  
|||||  
DB 55700 GACAAATGTTAAAGACAAACTTTGCACTAAGCGAACGTAAGCTGCCCCCATTTATCTC 55641  
|||||  
QY 148 148  
DB 55640 CTGTTCTAGTACCTGTGTGGAACGTGACATTCCTTACTACCCATACACCCCATCCGA 55581  
|||||  
QY 148 148  
DB 55580 GAAACCTCGATTATAGTCAGAGAGCTGTCTTCATTTCCTTGCTTAGGGACTTGG 55521  
|||||  
QY 148 148  
DB 55520 GTTTCCTCCACATTTTCTGTGAGCTTTCAGAAATTTGTGTCCCTTGACCTGAGATPACA 55461  
|||||  
QY 148 148  
DB 55460 GGTGAGAGACAGTGGGAGTGTTCATTTATTAAGCATATATATCTGCAGAAATGCTGCT 55401  
|||||  
QY 148 148  
DB 55400 TGGTGTGTTTGTGTACAGCTTGCTGGGAATTTTCTTGATACGCGAGATATCAGTA 55341  
|||||  
QY 149 149  
DB 55340 TGAGATGCGCATATTCCTTAAATGTTTGCATTTTGCATTTTGCAGATCTTGTGACCATGCCCTG 55281  
|||||  
QY 156 HisIleSerHisAspGluLeu 162  
|||||  
DB 55280 CACATATTCGATGATGAGCTGA 55260



RESULT 9  
AC073896/c 182350 bp DNA linear HTG 24-AUG-2002  
LOCUS Homo sapiens chromosome 12 clone RP11-977G19, WORKING DRAFT  
DEFINITION AC073896, 5 unordered pieces.  
ACCESSION AC073896  
VERSION AC073896.37 GI:22450359  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 182350)  
AUTHORS Muzny,D.M., Adams,C., Adio-Oluola,B., Alt-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayala,M., Banks,T.,  
Barber,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D.,  
Bouch,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Correll,J.C., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.E.,  
Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichteberg,O., Lieu,C., Liu,D., Liu,Y., Louissegh,H.,  
Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapa,P., Martin,R., Matindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Monabdat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunou,G.,  
Otagunye,N., Oyedro,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rolubokan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
Scherer,S., Scott,G., Shen,R., Shooshart,I., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tantony,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Taman,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.  
TITLE Unpublished  
JOURNAL Direct Submission  
REFERENCE 2 (bases 1 to 182350)  
AUTHORS Worley,K.C.  
JOURNAL Direct Submission  
TITLE Submitted (01-JUL-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 3 (bases 1 to 182350)  
AUTHORS Worley,K.C.  
JOURNAL Direct Submission  
TITLE Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
COMMENT On Aug 23, 2002 this sequence version replaced gi:22296903.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Drafting Center Code: BCM

Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: HBWI  
Center clone name: RP11-977G19  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Sequencing vector: M13;  
Chemistry: Dye-terminator Big Dye 85% of reads  
Chemistry: Dye-terminator Big Dye 85% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 179279 bases at least Q40  
Consensus quality: 180581 bases at least Q30  
Consensus quality: 181073 bases at least Q20  
Estimated insert size: 174258; sum-of-coverage estimation  
Quality coverage: 8x in Q20 bases; sum-of-coverage estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 5 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 2893: contig of 2893 bp in length  
\* 2894 2993: gap of unknown length  
\* 2994 6686: contig of 3693 bp in length  
\* 6687 6787: gap of unknown length  
\* 6787 17251: contig of 10465 bp in length  
\* 17252 17351: gap of unknown length  
\* 17352 85894: contig of 68543 bp in length  
\* 85895 85994: gap of unknown length  
\* 85995 182350: contig of 96356 bp in length.  
FEATURES  
source location/qualifiers  
1..182350  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-977G19"  
/chromosome="12"  
BASE COUNT 46453 a 45416 c 43855 g 46174 t 452 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,93e-37 Length: 182350  
Score: 481.00 Matches: 103  
Percent Similarity: 63.80% Conservative: 1  
Best local Similarity: 63.19% Mismatches: 6  
Query Match: 56.72% Indels: 53  
DB: Gaps: 1  
US-10-082-502-20 (1-162) x AC073896 (1-182350)  
QY 49 ValProtyrAlaArgSerGluAlaHisLeuThrGluLeuLeuGluGluIleCysAspArg 68  
|||||  
Db 172027 GTGGCTTATGCCCGCTCAGAGCCACCTCACAAGCTGCGAGAGATATGACCGG 171968  
QY 69 MetysGluYrGlyGluGluIleAspProSerThrHisArgLysAsnTyValArgVal 88  
|||||  
Db 171967 ATGAAGAGATGGGGAACAGATTGATCTTCCACCCATGCCAAGAACTACGTCGCTGA 171908  
QY 89 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGluGluGluGluIleArgLysSerAsp 108  
|||||  
Db 171907 GTGGCCGGAATGAGATGAGATGCACTGACCTACAGAGCAATCCGATGACTCAGAT 171848  
QY 109 ILeSerGlyThrLeuLysPheAla----- 116  
|||||  
Db 171847 ATTAGCGGACACCTCAAGTTGCGGTGAGCTATGAGATGAGTGTCTTGACAT 171788  
QY 116 ----- 116  
Db 171787 TAAAGGCTGTGGAGAACCAATTAGAAAGTTAGGCTGATATCTTCTCTCTTTGG 171728



OY 116 ----- 116  
DB 171727 AAGTGAAGCAGACGCTTATTCGCTGTCACCCATTTCCTGGTTCGACG 171668  
OY 117 CCGGUSGJIEVIGLUGLUTYrGLuApGLUeULeGLUePhSeSeArGLuA 136  
DB 171667 TGTGAGACATTGTGGAGATACGAGATGATGATTCATTCCTCCGAGAGGCT 171608  
OY 137 AspaVallyAspLyLeucCySeSerLySaRgThraP-LeucYsaSPHISAlaLeuH1 156  
DB 171607 GACATGTGTAAGACAACTTGTGAGTAGAGACAGTAGCTGCCCTTATTCCT 171548  
OY 156 sllaser 158  
DB 171547 CTGTTCT 171541  
RESULT 10 184762 bp DNA 1linear HTG 24-AUG-2002  
AC023500  
LOCUS Homo sapiens chromosome 12 clone RP11-183H16, WORKING DRAFT  
DEFINITION  
AC023500  
VERSION AC023500.32 GI:22095070  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 184762)  
Muzny,D.M., Adams,C., Adio-Oduola,B., All-oman,F.R., Allen,C.,  
Alshrocks,S.L., Amaralung,H.C., Are,J.R., Ayala,M., Banks,T.,  
Barberia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonini,D.,  
Bouch,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homel,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsdon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okwuon,G.,  
Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peary,J., Peters,L.,  
Peters,L., Pickens,R., Primus,E., Rofle,M., Ruiz,S., Savary,G.,  
Rivers,R., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vaequez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wlecyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE2 (bases 1 to 184762)  
Unpublished  
Direct Submission

AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Mortley, K.C.  
Direct Submission  
Submitted (15-FEB-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 184762)  
Mortley, K.C.  
Direct Submission  
Submitted (24-AUG-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Aug 4, 2002 this sequence version replaced gi:22094227.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Drafting Center Code: BCM  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: HARC  
Center clone name: RP11-183H16  
----- Summary Statistics  
Sequencing vector: Plasmid  
Sequencing vector: M13  
Chemistry: Dye-terminator Big Dye 988 of reads  
Chemistry: Dye-terminator Big Dye 988 of reads  
Assembly program: Phrap version 0.990329  
Consensus quality: 183845 bases at least Q40  
Consensus quality: 184849 bases at least Q30  
Consensus quality: 185447 bases at least Q20  
Estimated insert size: 159026; sum-of-contigs estimation  
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 9 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 2040: contig of 2040 bp in length  
\* 2041 2140: gap of unknown length  
\* 2141 4170: contig of 2030 bp in length  
\* 4171 4270: gap of unknown length  
\* 4271 8294: contig of 4024 bp in length  
\* 8295 8394: gap of unknown length  
\* 8395 14196: contig of 5802 bp in length  
\* 14197 14296: gap of unknown length  
\* 14297 33076: contig of 18780 bp in length  
\* 33077 33176: gap of unknown length  
\* 33177 52295: contig of 1919 bp in length  
\* 52296 52395: gap of unknown length  
\* 52396 84185: contig of 31790 bp in length  
\* 84186 84285: gap of unknown length  
\* 84286 120023: contig of 35738 bp in length  
\* 120024 184762: gap of unknown length  
\* 120124 184762: contig of 66639 bp in length.

FEATURES  
SOURCE

1. 184762  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-183H16"  
BASE COUNT 47425 a 44751 c 45101 g 46673 t 812 others  
ORIGIN

Alignment Scores:  
Pred. No.: 5.01e-37 Length: 184762  
Score: 481.00 Matches: 103  
Percent Similarity: 63.80% Conservative: 1

Best Local Similarity: 63.19% Mismatches: 5  
Query Match: 56.72% Indels: 53  
DB: 2 Gaps: 1

US-10-082-502-20 (1-162) x AC023500 (1-184762)

OY 49 ValProtyrAlaAysSerGluAlaHisIeuThrGluLeuLeuGluGluIleCysAspArg 68  
|||||  
DB 43486 GTGCTTATGCCCCCTCAGAGCCACCTCAGACAGCTGCGAGAGATATGTGACCGG 43545

OY 69 MetLysGluTyrgLysGluGlnIleAspProSerThrHisArgLysAsnTyValArgVal 88  
|||||  
DB 43546 ATGAAGAGGATGGGGAACATATGATCTTCACCCATCCCAAGAACTAGTACTGTA 43605

OY 89 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgLysAspSerAsp 108  
|||||  
DB 43606 GTGGCCCGGAATGGAGATCCAGTCAACTGAGCTACAAAGCAATCCGAATGACTCAGAT 43665

OY 109 IleSerGlyThrLeuLysPheAla----- 116  
|||||  
DB 43666 ATTAGCGGACCCCTCAAGTTTGGCGTGAAGCTATGGAATCGGTCTTGACAT 43725

OY 116 ----- 116  
DB 43726 TAAGGGTTGTGGAGAACCAATTAGAAAGTTAGGCTGATATCTCTGCTCTCTTTGG 43785

OY 116 ----- 116  
DB 43786 AAGTGAAGGCAAGAGCCCTTATATGCCCTGTGTACCCCATTTCTCCCTGGTTGGCAG 43845

OY 117 CysGluSerIleValGluGluTyrgLysAspGluLeuIleGluPhePheSerArgGluAla 136  
|||||  
DB 43846 TGTAGAGCATGTGGAGGAATGAGAGATGAACTCATGTAATCTTTCCCGAGAGGCT 43905

OY 137 AspAsnValLysAspLysLeuCySerLysArgThrAsp-LeuCyAspHisAlaLeuH1 156  
|||||  
DB 43906 GACAATTTAAAGCAAACTTTCAGTAAGCAAGTAAGTACGCCCCCATTTATCTC 43965

OY 156 sIleSer 158  
|||  
DB 43966 CTGTCTT 43972

RESULT 11  
AC090489 207424 bp DNA linear ROD 26-JAN-2002  
LOCUS AC090489  
DEFINITION Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence.  
AC090489  
VERSION AC090489.8 GI:18376849  
KEYWORDS HTG.  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 207424)  
REFERENCE  
AUTHORS McCormick, W.R., de la Bastide, M., Spiegel, L., Preston, R.,  
Kirchoff, K., Kuit, K., Nascimento, L., Zucavero, T., Ballia, V.,  
Bell, M., Baker, J., Santos, L., Miller, B., Katzenberger, F.,  
Muller, S., King, L., Yang, C., Palmer, L., O'Shaughnessy, A. and  
Dedhia, N.  
TITLE Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 207424)  
AUTHORS McCormick, W.R.  
TITLE Direct Submission  
JOURNAL Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
3 (bases 1 to 207424)  
REFERENCE  
AUTHORS McCormick, W.R.  
TITLE Direct Submission  
JOURNAL Submitted (26-JAN-2002) Lita Annenberg Hazen Genome Sequencing

Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
On Jan 26, 2002 this sequence version replaced gi:18201765.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest.

FEATURES  
source  
1..207424  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-104010"  
/clone\_1lb="RPCI-23"  
48132..48165  
misc\_feature  
/note="We believe the assembly to be correct. The sequence  
is a mononucleotide (T) repeat in which the exact number  
of Ts is unknown. Other subclones in the region show one  
fewer T than that represented in the assembly."

BASE COUNT 52613 a 48671 c 49663 g 56477 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.83e-33 Length: 207424  
Score: 446.50 Matches: 98  
Percent Similarity: 48.77% Conservative: 1  
Best Local Similarity: 48.28% Mismatches: 1  
Query Match: 52.65% Indels: 103  
DB: 10 Gaps: 1

US-10-082-502-20 (1-162) x AC090489 (1-207424)

OY 49 ValProtyrAlaAysSerGluAlaHisIeuThrGluLeuLeuGluGluIleCysAspArg 68  
|||||  
DB 75789 GTACCTTATGCCCCCTCAGAGCCACCTCAGACAGCTGCTGAGAGGTGTGACCGA 75848

OY 69 MetLysGluTyrgLysGluGlnIleAspProSerThrHisArgLysAsnTyValArgVal 88  
|||||  
DB 75849 ATGAAGAGGATGGGGAACATATGACCTTCACCCACCGCAAGAACTAGTACGGCTC 75908

OY 89 ValGlyArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgLysAspSerAsp 108  
|||||  
DB 75909 GTGAGCCGGAATGGAGATGAGATGAACTTACAGGCGATCCGAATGACTCAGAT 75968

OY 109 IleSerGlyThrLeuLysPheAla----- 116  
|||||  
DB 75969 ATCAGCGGACCCCTCAAGTTTGGCGTGAAGTACCGCTGTGAGTACTGTTCTTG 76028

OY 116 ----- 116  
DB 76029 GAAGTATAGGATTCCTCGGACAGAGATGAGATGCGAAGTGGTTCCTAC 76088

OY 116 ----- 116  
DB 76089 TTGTAACCTAAGACACTCAGGGGAGCGGTTGAAGATCAGCTTAAGTTGAGCGTGGC 76148

OY 116 ----- 116  
DB 76149 CTAGTCAATAGTACGTTGAGAGCCAGCTGGGACAGATATAGAGCCCTGTATAAAT 76208

OY 116 ----- 116  
DB 76209 CCCCACCCCAACCAACAGTAGAGAGCTGAGCAGTCTTGGCCACCTTGGGA 76268

OY 117 ----- 117  
DB 76269 TCTGGTCAATTCCTGTCTTCTTGGTCCAGAGTGAAGACATTTGGAAGATACGAG 76328

OY 126 AspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLysAspLysLeuCySer 145  
|||||  
DB 76329 GATGAGCTTATCGAATCTTCTCCAGAGGCTGACAACTTAAGCAAACTTTGCACT 76388

OY 146 LysArgThr 148  
 Db 76389 AAGCGACA 76397  
 RESULT 12  
 AC109891  
 LOCUS AC109891 101666 bp DNA linear HTG 13-JUL-2002  
 DEFINITION Rattus norvegicus clone CH230-319E6, \*\*\* SEQUENCING IN PROGRESS.  
 \*\*\* 47 unordered pieces.  
 ACCESSION AC109891  
 VERSION AC109891.3 GI:21738196  
 KEYWORDS HTG: HTGS PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Ratus.  
 1 (bases 1 to 101666)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,  
 Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbarta,J., Benton,J., Blmage,K., Blankenburg,K., Bonin,D.,  
 Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Donthwite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,U., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
 Homel,F., Howard,S., Huber,J., Hulky,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudh,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Kovah,J., Kovar,C.,  
 Kirovsky,C.J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louisedge,H.,  
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Medori,M., Mei,G., Metzger,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokoko,S., Oguh,M., Okunolu,G.,  
 Orduyue,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Qulles,M., Ren,Y.,  
 Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
 Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Usmani,K., Vasquez,L., Vera,V., Willison,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G. and Gldbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 101666)  
 Worley,K.C.  
 Direct Submission  
 Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 101666)  
 Worley,K.C.  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Jul 12, 2002 this sequence version replaced gi:18660222.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GSKY  
 Center clone name: CH230-319E6  
 ----- Summary Statistics  
 Sequencing vector: Plasmid:  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 60229 bases at least Q40  
 Consensus quality: 63959 bases at least Q30  
 Consensus quality: 66730 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 47 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1 1423: contig of 1423 bp in length  
 \* 1424 1523: gap of unknown length  
 \* 1524 2972: contig of 1449 bp in length  
 \* 2973 3073: gap of unknown length  
 \* 3073 4717: contig of 1645 bp in length  
 \* 4718 4817: gap of unknown length  
 \* 4818 6532: contig of 1815 bp in length  
 \* 6533 6732: gap of unknown length  
 \* 6733 8041: contig of 1308 bp in length  
 \* 8041 8141: gap of unknown length  
 \* 8141 10043: contig of 1903 bp in length  
 \* 10044 10143: gap of unknown length  
 \* 10144 11237: contig of 1094 bp in length  
 \* 11238 11337: gap of unknown length  
 \* 11338 12362: contig of 1025 bp in length  
 \* 12363 12462: gap of unknown length  
 \* 12463 14100: contig of 1638 bp in length  
 \* 14101 14201: gap of unknown length  
 \* 14201 15840: contig of 1640 bp in length  
 \* 15841 15940: gap of unknown length  
 \* 15941 17053: contig of 1113 bp in length  
 \* 17054 17153: gap of unknown length  
 \* 17154 18836: contig of 1683 bp in length  
 \* 18837 18936: gap of unknown length  
 \* 18937 20137: contig of 1201 bp in length  
 \* 20138 20237: gap of unknown length  
 \* 20238 21736: contig of 1439 bp in length  
 \* 21737 21836: gap of unknown length  
 \* 21837 23619: contig of 1783 bp in length  
 \* 23620 23719: gap of unknown length  
 \* 23720 25907: contig of 2188 bp in length  
 \* 25908 26007: gap of unknown length  
 \* 26008 27406: contig of 1399 bp in length  
 \* 27407 27506: gap of unknown length  
 \* 27507 28837: contig of 1321 bp in length  
 \* 28838 28937: gap of unknown length  
 \* 28938 30229: contig of 1282 bp in length  
 \* 30230 30329: gap of unknown length  
 \* 30330 32012: contig of 1683 bp in length  
 \* 32013 32112: gap of unknown length  
 \* 32113 33198: contig of 1086 bp in length  
 \* 33199 33298: gap of unknown length  
 \* 33299 35519: contig of 2221 bp in length  
 \* 35520 35619: gap of unknown length  
 \* 35620 37677: contig of 2058 bp in length  
 \* 37678 37777: gap of unknown length



BASE COUNT 175 a 150 c 171 g 149 t 4 others

## ALIGNMENT SCORES:

Pred. No.: 2,73e-17 Length: 649  
Score: 257.00 Matches: 49  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 30.31% Indels: 0  
DB: Gaps: 9

US-10-082-502-20 (1-162) x AF186113 (1-649)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20

DB 164 CGGAGAGACCCAGATCTCCAGTGTGAGACATGACGGCTGTGTGGATGATGAATGATG 223

OY 21 GIU11eAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40

DB 224 GAAATTCGCCAGGTGAGACCCCAAGAACATTCAGATGGATCTTTCCGATCAATCCA 283

OY 41 AspGlySerGlnSerValValGluVal 49

DB 284 GATGGACCCAGTCAGTGTGGAGGTA 310

RESULT 14 BC001027 824 bp mRNA linear PRI 12-JUL-2001

DEFINITION Homo sapiens, transmembrane protein 4, clone MGC:1545

IMAGE:3344786, mRNA, complete cds.

ACCESSION BC001027.1 GI:12654402

VERSION BC001027.1

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (17-NOV-2000) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Galtherburg, Maryland;

Web site: http://www.nisc.nih.gov/

Contact: nisc.mgc@nih.gov

Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,

Benjamin, B., Blakeley, R.W., Boufara, G.G., Brinkley, C., Brooks, S.,

Dierlich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,

Lim, M., Maduco, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,

McDowell, E., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J.,

Thompson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,

Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

Series: RAL Plate: 5 Row: P Column: 13

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 6014631.

Location/Qualifiers

1. 824

/organism="Homo sapiens"

/db\_xref="locusID:10330"

/db\_xref="taxon:9606"

## CDS

BASE COUNT 221 a 194 c 233 g 176 t

ORIGIN

Alignment Scores:

Pred. No.: 3.65e-17 Length: 824

Score: 257.00 Matches: 49

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 30.31% Indels: 0

DB: Gaps: 9

US-10-082-502-20 (1-162) x BC001027 (1-824)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20

DB 310 CGGAGAGACCCAGATCTCCAGTGTGAGACATGACGGCTGTGTGGATGATGAATGATG 369

OY 21 GIU11eAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40

DB 370 GAAATTCGCCAGGTGAGACCCCAAGAACATTCAGATGGATCTTTCCGATCAATCCA 429

OY 41 AspGlySerGlnSerValValGluVal 49

DB 430 GATGGACCCAGTCAGTGTGGAGGTA 456

RESULT 15 AK094445 2378 bp mRNA linear PRI 15-JUL-2002

LOCUS AK094445

DEFINITION Homo sapiens CDNA FLJ37126 fls, clone BRACE2022693, weakly similar

to transmembrane protein 4.

ACCESSION AK094445.1 GI:21753508

VERSION AK094445.1

KEYWORDS Oligo capping; fls (full insert sequence).

SOURCE Homo sapiens cerebellum CDNA to mRNA, clone\_11b:BRACE2

clone:BRACE2022693.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,

Otsuki, T., Sato, H., Makamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,

Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,

Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,

Magatsuna, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,

Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K.,

and Isogai, T.

NEO human cDNA sequencing project

Unpublished

2 (bases 1 to 2378)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-Kamitani, Kisarazu, Chiba 292-0812, Japan

(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; CDNA full insert sequencing:

Research Association for Biotechnology (RAB); CDNA library

Key Technology Center etc.); 5 - 6 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

FEATURES RAB; annotation: HRI and RAB.  
Location/Qualifiers  
Source 1..2378

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="BRACE2022693"  
/tissue\_type="cerebellum"  
/clone\_lib="BRACE2"  
/note="cloning vector: PME18SFL3"

BASE COUNT 723 a 432 c 477 g 746 t  
ORIGIN

Alignment Scores:

Pred. No.:	1.73e-15	Length:	2378
Score:	246.00	Matches:	49
Percent Similarity:	62.30%	Conservative:	27
Best Local Similarity:	40.16%	Mismatches:	44
Query Match:	29.01%	Indels:	2
DB:	9	Gaps:	1

US-10-082-502-20 (1-162) x AK094445 (1-2378)

```

OY 41 AspGlySerGlnSerValValGluValProtyrAlaArgSerGluValHisLeuThrGlu 60
    |||||
Db 80 GACGACTGACTACTACAGGACGAGATCCCTACCTCAGTCGAGCGCTTCTACCGGAT 139
    |||||

OY 61 LeuGluGluIleuLeuCyAspArgMetIleGluIleGluGluGluGluGluGluGlu 80
    |||||
Db 140 CTTTGGAGAAAGCTGTGAGCGAATGACGAACTACAAAGCTTGAAGAACCTGTGACG 199
    |||||

OY 81 HisArgIleAsnIleValArgValValGluValGluValGluValGluValGluVal 100
    |||||
Db 200 AAGGAGAGAACTTTCAGAGATTCGCTCTGAGAAAGAGACAAATATACCAAGAAATT 259
    |||||

OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuIlePheAlaCysGluSerIle 120
    |||||
Db 260 AAAAATTTGATTTTATTCGATGCTTACAGACCTTGAATTTGCGTGTGAACATATA 319
    |||||

OY 121 ValGluIleuTyrgluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
    |||||
Db 320 ATAGAGAGATGATGAAATGAAATATCCTCACTTATCGCCGAGAGACACACTATAGCT 379
    |||||

OY 141 AspIleuGlySerIleuArgIleThrAspLeuGlyAspHisAlaLeuHisIleSerHisAsp 160
    |||||
Db 380 GACAAAGCTGTGACGTAAGAAATCAGATCTGTGAAACTTCT-----GCTAATCATACT 433
    |||||

OY 161 GluLeu 162
    |||||
Db 434 GAGCTC 439
    |||||

```

Search completed: March 31, 2003, 03:42:06  
Job time : 1675.7 secs

- 1: em\_eestba:\*
- 2: em\_eesthu:\*
- 3: em\_eestln:\*
- 4: em\_eestlov:\*
- 5: em\_eestlov:\*
- 6: em\_eestpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_eesti:\*
- 10: gb\_estl2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estlun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_huv:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_man:\*
- 24: em\_gss\_mam:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

1	848	100.0	588	14	BM832828
2	848	100.0	591	14	BM832868
3	848	100.0	611	14	BM712898 K-ESTO10
4	848	100.0	613	14	B1712899 I948F09.
5	848	100.0	697	12	BO686530 AGENCOUR
6	848	100.0	719	12	BF4464334
7	848	100.0	742	12	BE733900
8	848	100.0	745	12	BE732449
9	848	100.0	930	14	BE791763
10	848	100.0	1117	14	BM437660
C 11	843	99.4	589	13	BM552907
C 12	842	99.3	590	13	BI791489
C 13	840	99.1	783	13	BO421827
C 14	835	98.5	868	14	BI550436
C 15	833	98.2	791	12	BO421069
C 16	830	97.9	665	10	BS6704443
C 17	830	97.9	674	14	AM9913565
C 18	830	97.9	754	14	BO202181
C 19	830	97.9	754	14	BO202181
C 20	830	97.9	772	11	AK013568
C 21	830	97.9	808	13	AK007914
C 22	830	97.9	843	13	BI408525
C 23	830	97.9	853	11	BI413539
C 24	830	97.9	1255	11	AK013014
C 25	827	97.5	773	13	AK019927
C 26	825	97.3	657	14	BI550162
C 27	822	96.9	818	12	BO211867
C 28	822	96.9	818	12	BE870315
C 29	821	96.8	887	13	BI763950
C 30	818	96.5	802	12	BI766759
C 31	817	96.3	678	14	BE608664
C 32	817	96.3	891	13	BE309953
C 33	816	96.2	709	13	BI408381
C 34	815	96.1	549	10	BI411532
C 35	815	96.1	593	10	BE198865
C 36	815	96.1	720	13	BI762927
C 37	809	95.4	797	13	BI909759
C 38	808	95.3	783	13	BI763984
C 39	808	95.3	914	12	BE796495
C 40	807	95.2	746	13	BI551335
C 41	799	94.2	765	12	BF313281
C 42	794	93.6	475	10	AW414257
C 43	794	93.6	856	12	BF965088
C 44	792	93.4	595	10	AW414039
C 45	792	93.4	856	13	BI454180

RESULT	1
BM832828	
LOCUS	568 bp, mRNA
DEFINITION	K-ESR01L7375.S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-11-B03
ACCESSION	BM832828
VERSION	BM832828.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 588)  
Kim,S., Hahn,Y., Oh,T.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and



TITLE  
JOURNAL  
COMMENT

Kim Y.S.  
21C Frontler Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 11 row: B column: 03  
High quality sequence stop: 588.  
Location/Qualifiers

FEATURES  
source

1..588  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SSNU484s1-11-B03"  
/clone\_1lb="SSNU484s1"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site:1: EcoRI Site:2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 Promotor as 5' primer and M(DT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."

BASE COUNT 149 a 140 c 176 g 123 t

ORIGIN

Alignment Scores:  
Pred. No.: 6,23e-103 Length: 588  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-20 (1-162) x BM832828 (1-588)

QY 1 ATGATGSeGlnAspLeuHtScyGlyAlaCyAArgAlaLeuValAspGluLeuGluTTP 20  
DB 62 CGAGAGACCGAGATCTCCACGTGTGAGCATGACGGCTCTGTCGATCAACTAGATCG 121  
QY 21 GtUleAlaGlnValAspProLyStyTrtIleGlnMetGlySerPheArgTtleAsnPro 40  
DB 122 GAAATTGCCAGGTGAGCCCAAGAGACCAATTCAGATGGATCTTCCGGATCAATCCA 181  
QY 41 AspGlySerGlnSerValValGluValProTyAlaArgSerGlnAlaHisLeuThrGlu 60  
DB 182 GATGGACGCACTGATGTGTGAGGTGCTTATGCCCGCTCAGAGGCCACCTCACAGAG 241

QY 61 LeuLeuGluGluIleCysAspArgMetLyGluTyGlyGluGlnIleAspProSerThr 80  
DB 242 CTCCTGAGCAGATATCTGACCCGATCAAGAGACTATGGGACATTCATCTCCACC 301  
QY 81 HisArgLyAspTrpValArgValValGlyArgAsnGlyGluSerGluLeuAspLeu 100  
DB 302 CATCGCAAGAACTACGTACTGTGTGGCCGGAATGGAAATCCAGTCACTGAGACCTA 361  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLySphaeIacysGluSerIle 120  
DB 362 CAAGGATCGCATGACATCCAGATATTAGGCGACCCCTCAAGTTTGGTGTGAGACATT 421  
QY 121 ValGluGluTyGlyGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLy 140  
DB 422 GTGAGAGAAATACAGAGATGAACATCATTTCTTCCGAGAGGCTGCACAAATGTTAA 481  
QY 141 AspLyLeuCySerLyAspArgTrpAspLeuCyAspHisAlaLeuHisIleSerHisAsp 160  
DB 482 GACAAACTTTCAGTAAAGCAACAGATCTTTGTGACCAATGCCCTGCACATATGCATGAT 541  
QY 161 GluLeu 162  
DB 542 GAGCTA 547

RESULT 2  
BM832888 591 bp mRNA linear EST 06-MAR-2002  
LOCUS K-EST0107445 SSNU484s1 Homo sapiens cDNA clone SSNU484s1-11-H01  
DEFINITION 5', mRNA sequence.  
ACCESSION BM832888  
VERSION BM832888.1 GI:19189297  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 591)  
Kim Y.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.  
21C Frontler Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 11 row: H column: 01  
High quality sequence stop: 591.  
Location/Qualifiers

FEATURES  
source

1..591  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SSNU484s1-11-H01"  
/clone\_1lb="SSNU484s1"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site:1: EcoRI; Site:2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The cDNA vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of

competent cells E. coli Top10F' by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(NT)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method.

BASE COUNT 150 a 142 c 176 g 123 t  
ORIGIN

## Alignment Scores:

Pred. No.: 6,28e-103 Length: 591  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-20 (1-162) x BM832888 (1-591)

OY 1 ArgArgSerGlnAspLeuHsCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
Db 62 CGAGAGAGCCAGCATCTCCAGTGTGAGCATCAGGCGCTGTGTGATGACATGATGATG 121  
OY 21 GlnTleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
Db 122 GAAATGCGCCAGGTGAGCAACCAAGACCATTCAGATGATGATCTTCCGATCAATCCA 181  
OY 41 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
Db 182 GATGGAGGACCATCGATGTTGAGGTGCGCTTATGCCCGCTCAGAGCCACCTCAGAG 241  
OY 61 LeuLeuGluGlnIleCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 80  
Db 242 CTGCTGGAGGAGATATGTACCGGATGAGAGATGAGGAGAACATGATGATCTTCCACC 301  
OY 81 HisArgLysAsnTyrValArgValAlaGluArgAsnGlyLysSerSerGluLeuAspLeu 100  
Db 302 CATCCAGAGACATACGTACGTAGTGGCGGAGATGAGATCCAGTCAATGACCTA 361  
OY 101 GlnGlyIleArgGlnLeaSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
Db 362 CAGGCAATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421  
OY 121 ValGlnGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
Db 422 GTGGAGGAATACGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 481  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisLeuSerHisAsp 160  
Db 482 GACAACTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541  
OY 161 GluLeu 162  
Db 542 GAGCTA 547

## RESULT 3

LOCUS B1712899

DEFINITION B1712899 611 bp mRNA linear EST 11-MAR-2002

1d98f09.y1 HR85 islet Homo sapiens cDNA clone IMAGE:508582.5'

sequence: similar to TR:O9Y2B0 O9Y2B0 Type II MEMBRANE PROTEIN. ; mRNA

ACCESSION B1712899

VERSION B1712899

KEYWORDS EST. GI:15688594

SOURCE human.

## ORGANISM

## REFERENCE

## AUTHORS

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 611)  
Melton, D., Brown, J., Kenty, G., Penmutt, A., Lee, C., Keestner, K.,  
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, J., Cardenas  
, M., Gibbons, M., McCann, R., Cole, R., Tsagarelis, R., Williams, T.,  
Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohpc.harvard.edu

## JOURNAL

## COMMENT

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
Seq primer: -40RP from GIBCO  
High quality sequence stop: 463.

## FEATURES

## source

1. 611  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5085832"  
/clone\_1bp="HR85 islet"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
Site-selected on agarose gel. Average insert size  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue,  
Metabolism Div. (Alan Pernutt Lab), Washington University,  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

BASE COUNT 157 a 143 c 184 g 127 t  
ORIGIN

## Alignment Scores:

Pred. No.: 6.6e-103 Length: 611  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-20 (1-162) x B1712899 (1-611)

OY 1 ArgArgSerGlnAspLeuHsCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
Db 90 CGAGAGAGCCAGCATCTCCAGTGTGAGCATCAGGCGCTGTGTGATGACATGATGATG 149  
OY 21 GlnTleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
Db 150 GAAATGCGCCAGGTGAGCAACCAAGACCATTCAGATGATGATCTTCCGATCAATCCA 209  
OY 41 AspGlySerGlnSerValValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
Db 210 GATGGAGGACCATCGATGTTGAGGTGCGCTTATGCCCGCTCAGAGCCACCTCAGAG 269  
OY 61 LeuLeuGluGlnIleCysAspArgMetLysGluTyrGlyGlnIleAspProSerThr 80  
Db 270 CTGCTGGAGGAGATATGTACCGGATGAGAGATGAGGAGAACATGATGATCTTCCACC 329  
OY 81 HisArgLysAsnTyrValArgValAlaGluArgAsnGlyLysSerSerGluLeuAspLeu 100

Db	330	CATCCCAAGAACTACGTACTGTCAGTGGGCCGGAATGAGAATCCAGTCACTGACCGACCTA	389
Oy	101	GInglyIleArgYIIAspSerAspIleSerGlyThrIleuAspPhalaICysglusertlle	120
Db	390	CAAGGCATCCGAATGACTCAGATATTAGCGCACCCCTCAAGTTTGCGGTGAGAGCAATT	449
Oy	121	VAlGlugluTyrgLUAspGLuLeuIllegluPheSerArgGLUALAspaSnyVallys	140
Db	450	GTGGAGGAATAAGAGATGAACTCATGTGAATCTTTCCCGAGAGCGTAGACAATGTMAA	509
Oy	141	AAspyLSLeuCysSerLyArGThrAspLeuCySaaphISalaleuHIsIleserHisasp	160
Db	510	GACCAACTTTCGATAGAGGAACAGATCTGTGTGACATGCCCTGCACATATCGCATGAT	569
Oy	161	GIuleu 162 	
Db	570	GAGCTA 575	
RESULT 4			
B0668530			
LOCUS	B0668530	673 bp	mRNA linear EST 15-JUL-2002
DEFINITION	AGENCOURT_82111118 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:6276413		
ACCESSION	5', mRNA sequence.		
VERSION	B0668530		
KEYWORDS	B0668530.1 GI:21778777		
SOURCE	EST.		
ORGANISM	human.		
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 673)		
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgrpbs-r@mail.nih.gov Tissue Procurement: ATCC		
	CDNA Library Preparation: Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: L1CM2461 row: c column: 06		
	High quality sequence start: 208		
	High quality sequence stop: 350.		
FEATURES	Location/Qualifiers		
source	1..673		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6276413"		
	/clone_lib="NIH-MGC_102"		
	/tissue_type="epidermoid carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: salivary gland; Vector: pOMV7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming, directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH_MGC Library."		
BASE COUNT	196 a 149 c 184 g 144 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	7.62e-103	Length:	673
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

QY	1	ArgAnSerGlnAspLeuHisGlySerLysAlaCysArgAlaLeuValAspGluLeuGluTrp	20
Db	47	CGAGAGAGCGAGAACTCTCCATGTGTGACATATGCAGGGCTGTGGTGAAGAACTGAAAGG	106
QY	21	GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro	40
Db	107	GAAATGTCCCAAGCTGAGCCCAAGAAAGACATTCAGATGGGATCTTTCCGGATCAATCA	166
QY	41	AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu	60
Db	167	GATGCGACCCAGTCACTGTGTGGAGCTGCTTATGCCCCCTCAGAGGCCACCTCAGAG	226
QY	61	LeuLeuGluGluIleCysAspArgMetLysGluTrpGluGluGlnIleAspProSerThr	80
Db	227	CTGCTGGAGGAAATGTGTGACCGGATGAAGGATGTGGGAAACAGATGATCTTCCACC	286
QY	81	HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyLysSerSerGlyLeuAspLeu	100
Db	287	CATGCCAAGAACTACGTACGTGTACGTGGCGCGGAATGGAAATCCAGGAACTGGACCTA	346
QY	101	GlnGlyIleLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
Db	347	CAAGCAGTCCGAAATGCAGTCAAGATTTGTAGCGGACCCCTCAAGATTGCTGTGACAGCAT	406
QY	121	ValGlnGluTrpGluAspGluLeuIleGluPhePheSerArgGluAlaAspArgValLys	140
Db	407	GTGGAGGAATTCAGAGAGTGAACCTCATTAATCTTTTCCGAGAGGCTGACAAATGTAA	466
QY	141	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp	160
Db	467	GACAAACCTTGGAGTAAAGCAACAGATCTTTGTGTGACATGGCCCTGCACATTCGATAT	526
QY	161	GluLeu 162	
Db	527	GAGCTA 532	
RESULT 5			
LOCUS	BF344334		
DEFINITION	602014708b1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4150512		
ACCESSION	BF344334		
VERSION	BF344334.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 697)		
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@b6femail.nih.gov		
	Tissue Procurement: David N. Louis, M.D.		
	CDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMNL at:		
	http://image.llnl.gov		
	Plate: LAM9413 row: p column: 01		
	High quality sequence stop: 692.		
FEATURES			
source	location/Qualifiers		
	1..697		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4150512"		
	/cclone_id="NCI_CGAP_Brn64"		
	/tissue_type="glioblastoma with EGFR amplification"		

```

/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV_SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI-CGAP Library."
BASE COUNT      173 a      167 c      220 g      137 t
ORIGIN

```

Alignment Scores:	
Pred. No.:	8.02e-103
Score:	848.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	100.00%
DB:	12
Length:	697
Matches:	16
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

QY	1	ArgAspSerGlnAspLeuHisIleCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp	20
Db	152	CGAGAGACCCAGCATCTCCACTGTGGAGCATGAGAGGCTCTGTGTGATGAACATGAGATGG	211
QY	21	GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro	40
Db	212	GAATATGCCAGGTGGAGCCCAAGAAAGACCATTCAGATGGGACTTTCCGATCATCA	271
QY	41	AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu	60
Db	272	GATGGCAGCCAGTCAGATGGGGAGGGGCTTATGCCCCGCTCAAGAGCCCCACCTCACAGAG	331
QY	61	LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr	80
Db	332	CTGCTGTGAGAGATATGTGTGACCGATGAAGAGACTATGGGAGAACATTCCTTCCTCCAC	391
QY	81	HisArgLysAsnTyrValAlaArgValIleGlyArgAsnGlyGluSerSerGluLeuAspLeu	100
Db	392	CATCGCAAGACTACTAGCTAGTGAAGGGCCGGAATGGAGAAATCCAGTAACGTGACCTA	451
QY	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuIlePheAlaGlyGluSerIle	120
Db	452	CAAGGCATCCGAATGCAGTCAATATTTACGGCACCCCTCAAGTTTGGCGGTAGAGACATT	511
QY	121	ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaIleAspAsnValIys	140
Db	512	GTGGAGGAATTAGAGAGATCAACTATTCCTTTCCCGAGAGGCTGCACATGTATAA	571
QY	141	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisasp	160
Db	572	GACAAACCTTTCGACGTAACGAAACAATCTTGTGTGACCATGCCCTGCACATATTCGATCAT	631
QY	161	GluLeu 162	
Db			
Db	632	GAGCTA 637	

```

cdna Library Preparation: Ling Hong/Rubin Laboratory
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLMN at image.lln.gov
plate: ILICM537 row: k column: 18
High quality sequence stop: 710.
Location/Qualifiers
1..719

```

Alignment Scores:	
Pred. No.:	8, 4e-103
Score:	848.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	12
Gaps:	0

QY	1	ATGATGSEGGAAASPLEUENISCSYSGIYLAACYSATGATALEVAALASGLULENGLUTRP	20
Db	108	CGAGAGAGCCAGGATCTCCACTGTGAGACATGCACAGCGCTCTGCTGGATGAACTGAANTGG	16
QY	21	GUUTLEAGLIVAlAsPProLysLYsThrIleGlnMetGLysERphearGILeasNPro	40
Db	168	GAATTTGCCAGGTGAGACCCCAAGAAAGACCATTCAGATGAGGATCTTTCCGATCAATCCA	22
QY	41	ASPGYSERGLINSErVAIaIGLIVaLProTYrILAaYSErGLUaIALHisLEuThrGlu	60
Db	228	GATGGCAGCAGCATCGTGTGTGAGAGTGGCTTATGCCCGCTACAGAGCCACACTACAGAC	28
QY	61	LEuLeuGLUGLUIleCYsaSPaRgMeLysGLUTrYrGLYGLUGLIIleAsPProSErThr	80
Db	288	CTGGCTGGAGGAGATATGTGACCGCGATCAAGAGATGAGGGAGAACAGATTCCTTCCACC	34
QY	81	HISATGLYSANrTYrVAIaRgVAIaIGLYrGASnGLYUSerSERGLULEuAsPLEu	100
Db	348	CATGCCAAGAACTACGTACGTGTGTGTGGCGGGATGAGAGAACTCACTGAACCTGAGACTA	40
QY	101	GLNGLYIlearGLIleasPSerAsPILeserGLYrThrLeuLysPheALaCYSGIUSerILE	120
Db	408	CAAGGCATCCCAATCGACTCGATGATATAGCGGCGACCCCTCAAGTTTGGCTGTGAGACATT	46
QY	121	VALGLUGLUTrYrGLIuAsPGLULEuILEGLUPheSErnrGLUaIALaSPaSnVALLYs	140
Db	468	GTGGAGGAATrCGAGGATGAAGCACTATTGAATCTTTTCCGAGAGGCTGCACATGTGTTAA	52
QY	141	ASPLYSLeuCYSerILYsArGTrHrAsPLeuCYsaSPHISaIALeUHNISILeserHISaSP	160
Db	528	GACAAACTTGTGACATAGCGAAGAGATCTTTGTGACATATCCCTGCACATATGCAATGAT	58
QY	161	GLIuLeu 162	
Db	588	GAGCTA 593	

LOCUS	BG824549	742 bp	mRNA	linear	EST 22-MAY-2001
DEFINITION	6037228454F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4868100 5',				
ACCESSION	mRNA sequence.				
VERSION	BG824549				
KEYWORDS	BG824549.1	GI:14172136			
SOURCE	EST.				
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 742)				
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LCM1737 row: k column: 13 High quality sequence stop: 740.				
FEATURES	Location/Qualifiers				
source	1..742				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:4868100"				
	/clone_lib="NIH_MGC_15"				
	/tissue_type="adenocarcinoma cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/note="Organ: colon; Vector: pOT87; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adapter: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using 2AP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"				
BASE COUNT	180 a	184 c	233 g	144 t	1 others
ORIGIN					
Alignment Scores:					
Pred. NO.:	8.81e-103	Length:	742		
Score:	848.00	Matches:	162		
Percent Similarity:	100.00%	Conservative:	0		
Best local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	12	Gaps:	0		
US-10-082-502-20 (1-162) x BG824549 (1-742)					
Oy	1 ArgAGSerGlnAApLeuHIScysGlyIacysArgAlaIeUValAspGluLeuGluTrp 20				
Db	201 CGGAGGACCCAGGAGCTCCACCTGAGACCAATGCACGGCTGTGGATGAACTAGAAATGG 260				
Oy	21 GluIleAGlnValAspProLysIshThrIleGlnMetGlySerPhaArgIleAsnPro 40				
Db	261 GAAATGGCCAGAGTGGACCCCAAGAGACCAATTCAGATGGGATCTTCCGGATCAATCCA 320				
Oy	41 AspGlySerGlnSerValIaGluValaPrProTyrAlaArgSerGluAlaHisLeuThrGlu 60				
Db	321 GATGACACCAAGTCAGTGGTGGAGGTGCTTAATGCCCTCAGAGGCCACTCACAAG 380				
Oy	61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80				
Db	381 CTGCTGGAGGAGATATGTGCACCGGATCAAGGAGTATGGGAAACAGATTGATCTCCAC 440				
Oy	81 HisArgLysAsnTyrValaIrgValaGluIyrArgAsnGlyLysSerGluLeuAspLeu 100				

DB	441	CATCCAGAACTACGTACGTAGTGGGCCCGAATTGGAATGCAGTAAGTAATGCAGCA	500
OY	101	GInglyIleAtyIIeAspSeraspIleSerGIyThrLeuysPhAlAcScyluSerIle	120
DB	501	CAAGCACCACCGAATGCAGACTATTTAGCGCACCCCTTCACATGGTCGTCAGACAT	560
OY	121	VAlGluGLuTyrgLAspPoluLeuIlegIuphePheSerArgGLuAlAspAsnValys	140
DB	561	GTGGAGGAAATACGAGAGATGAACTCATTTGAATCTTTTTCCCGAGAGCTGCACAATGTTAAA	620
OY	141	AsplylsLeucysSerLyAsrgrThraspleucysAspRHisAlaleuHIsIleSerHisap	160
DB	621	GACAACTTGTCAGTAGAGACAGACATGCTTTGTGACACATGCCCTGCACATATGCATGAT	680
OY	161	Gluleu 162	
DB	681	GAGCTA 686	
RESULT	8		
Locus	BE791763	745 bp	mRNA linear EST 20-SEP-2000
DEFINITION	BE791763	60158158661 NIH_MGC_7 Homo sapiens CDNA clone IMAGE:3936141 5'	
ACCESSION	BE791763	mRNA sequence.	
VERSION	BE791763		
KEYWORDS	EST.	NIH-MGC http://mgc.ncl.nih.gov/	
SOURCE	human.	National Institutes of Health, Mammalian Gene Collection (MGC)	
ORGANISM	Homo sapiens	Contact: Robert Strausberg, Ph.D.	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Email: cgaabg@mail.nih.gov	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	Tissue Procurement: DCD/DMP	
TITLE	Unpublished (1999)	CDNA Library Preparation: Ling Hong/Rubin Laboratory	
JOURNAL		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	
COMMENT		DNA Sequencing by: Incyte Genomics, Inc.	
		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov	
		Plate: LINC79 row: k column: 22	
		High quality sequence stop: 739.	
FEATURES		Location/Qualifiers	
source		1..745	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:3936141"	
		/clone_lib="NIH_MGC_7"	
		/tissue_type="small cell carcinoma"	
		/cell_line="MGC3"	
		/lab_host="DH10B (phage-resistant)"	
		/note="Organ: lung; Vector: pORF7; Site: 1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." "	
BASE COUNT	162 a	180 c	237 g 146 t
ORIGIN			
Alignment Scores:			
Pred. No.:	8	866-103	length: 745
Score:	848.00		Matches: 162
Percent Similarity:	100.00%		Conservative: 0
Best Local Similarity:	100.00%		Mismatches: 0
Query Match:	100.00%		Indels: 0
DB:	12		Gaps: 0

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
|||||  
Db 198 CGAGGAGCCAGGATCTCCACTGTGGAGCATGCGGGCTGTGGTGAAGTAAGATGG 257  
QY 21 GlnUleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
|||||  
Db 258 GAATTTCCCAAGTGGAGCCCAAGAGACCAATTCAGATGGATCTTCCCGATCAATCCA 317  
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
|||||  
Db 318 GATGCGAGCCAGTCAGTGTGGAGGTGCTTATGCGCCGCTCAAGGCCCACTCAGAGAG 377  
QY 61 LeuLeuGluGlnIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
|||||  
Db 378 CTCTGGAGAGATATGTGACCGAGTGAAGAGATATGGGAGACAGATTGATCTTCCACC 437  
QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 100  
|||||  
Db 438 CATGCGAAGAACTACGTAGCTGTAGTGGCCGCAATGAGAAATCCAGTGAAGTGAAGCTTA 497  
QY 101 GlnGlyTleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
Db 498 CAAGGATCCGAAATCGACATTCAGATATTAGCGGACCCCTCAAGTTGCGTGAAGACAT 557  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGlnAlaAspAsnValLys 140  
|||||  
Db 558 GTGGAGAGATACAGATGATACATTTGATTTTCCCGAGAGCTGACATGTTAA 617  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
|||||  
Db 618 GACAAACTTTCAGTAAAGCAGACAGATCTTTGACCAATGCCCTGCACATATCGATGAT 677  
QY 161 GlnLeu 162  
|||||  
Db 678 GAGCTA 683

RESULT 9  
LOCUS BQ437660 930 bp mRNA linear EST 24-MAY-2002  
DEFINITION AGENCOURT\_7897085 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6159838  
5' mRNA sequence.  
ACCESSION BQ437660  
VERSION BQ437660.1 GI:21176736  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 930)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DMP  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LLM13509 row: 1 column: 23  
High quality sequence stop: 515.  
Location/Qualifiers  
1. 930  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6159838"  
/clone\_id="NIH\_MGC\_72"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NCI;"

BASE COUNT 244 a 221 c 257 g 208 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,23e-102 Length: 930  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-10-082-502-20 (1-162) x BQ437660 (1-930)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
|||||  
Db 10 CGAGGAGCCAGGATCTCCACTGTGGAGCATGCGGGCTGTGGTGAAGTAAGATGG 69  
QY 21 GlnUleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
|||||  
Db 70 GAATTTCCCAAGTGGAGCCCAAGAGACCAATTCAGATGGATCTTCCCGATCAATCCA 129  
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
|||||  
Db 130 GATGCGAGCCAGTCAGTGTGGAGGTGCTTATGCGCCGCTCAAGGCCCACTCAGAGAG 189  
QY 61 LeuLeuGluGlnIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
|||||  
Db 190 CTCTGGAGAGATATGTGACCGAGTGAAGAGATATGGGAGACAGATTCATCTTCCACC 249  
QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGluSerSerGluLeuAspLeu 100  
|||||  
Db 250 CATGCGAAGAACTACGTAGCTGTAGTGGCCGCAATGAGAAATCCAGTGAAGTGAAGCTTA 309  
QY 101 GlnGlyTleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
Db 310 CAAGGATCCGAAATCGACATTCAGATATTAGCGGACCCCTCAAGTTGCGTGAAGACAT 369  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGlnAlaAspAsnValLys 140  
|||||  
Db 370 GTGGAGAGATACAGATGATGACATCTTGTTCCTCCGAGAGCTGCATATGTTAA 429  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
|||||  
Db 430 GACAAACTTTCAGTAAAGCAGACAGATCTTTGACCAATGCCCTGCACATATCGATGAT 489  
QY 161 GlnLeu 162  
|||||  
Db 490 GAGCTA 495

RESULT 10  
LOCUS BM552907 1117 bp mRNA linear EST 20-FEB-2002  
DEFINITION AGENCOURT\_5572552 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5466942  
5' mRNA sequence.  
ACCESSION BM552907  
VERSION BM552907.1 GI:18791172  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1117)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCTD/DMP  
CDNA Library Preparation: Rudin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)



DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM1968 row: c column: 07  
 High quality sequence stop: 709.

## FEATURES

## SOURCE

Location/Qualifiers

1. 1117  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5466942"  
 /clone.lib="NIH\_MGC\_41"  
 /tissue\_type="amelanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAGGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using Zap-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

BASE COUNT 299 a 258 c 320 g 239 t 1 others

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.62e-102 Length: 1117  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-082-502-20 (1-162) x BM552907 (1-1117)

OY 1 ArgargSerGlnaspLeuHiscysGlyAlaCysArgAlaLeuValaspGluLeuGluTTP 20  
 |||||||  
 DB 208 CGAGAGAGCCAGAGATCTCCACTGTGAGCATGAGGCGCTCTGTGATGACATGAGATGG 267  
 |||||||  
 OY 21 GlllIleAlaGlnValaspProlyslsYthrIleGlnMetGlySerPheargIleasPro 40  
 |||||||  
 DB 268 GAAATGGCCAGCTGGAGCCCAAGAACACATTCAGATGGGATCTTCCGATCAATCCA 327  
 |||||||  
 OY 41 AspGlySerGlnSerValAlaGluValProtyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 |||||||  
 DB 328 GATGGAGCCAGTCACTGAGTGGAGGCTTATGCGCCCTCAAGAGCCACCTCACAGAG 387  
 |||||||  
 OY 61 LeuLeuGluGluIleCysaspArgMetLysGluTyrGlyGluGlnIleasProSerThr 80  
 |||||||  
 DB 388 CTCCTGGAGAGATATGTACCCGATGAAGGAGATGGGGAACAGTTGATCTTCCACC 447  
 |||||||  
 OY 81 HisArgLysAsnTyrValaArgValAlaGlyArgasnGlyLuserSerGluLeuAspLeu 100  
 |||||||  
 DB 448 CATGCCAAGACTACCTAGCTAGTGGGCGGCAATGAGAAATCCAGTCACTGAGACTA 507  
 |||||||  
 OY 101 GlnGlyIleArgIleasPseraspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||||  
 DB 508 CAAGGATCCGATCGACATGATATTACGGCACCTCAAGTTTCCGTTGAGACAT 567  
 |||||||  
 OY 121 ValGluGluTyrGluaspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 |||||||  
 DB 568 GTGGAGAGATACGAGATGAATCTTCTTCCGAGAGAGCTGACATGTTAAA 627  
 |||||||  
 OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 |||||||  
 DB 628 GACAAACCTTTCAGATGAAGGAGACAGATCTTTGTGACACATGCCCTGCACATATCGCATGAT 687  
 |||||||  
 OY 161 GluLeu 162  
 |||||||  
 DB 688 GAGCTA 693  
 |||||||

RESULT 11  
 B1791489/c

LOCUS B1791489 599 bp mRNA linear EST 11-MAR-2002

DEFINITION Id98f09.x1 HR85 Islet Homo sapiens cDNA clone IMAGE:5085832 3'

similar to TR:Q9Y2B0 Q9Y2B0 TYPE II MEMBRANE PROTEIN. ; mRNA

sequence.

## ACCESSION

B1791489

## VERSION

B1791489.1 GI:15819214

## KEYWORDS

EST.

## SOURCE

human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 599)

## AUTHORS

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,  
 Lemishka,I., Scease,M., Brestelli,J., Gradwohl,G., Clifton,S.,  
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistein,A.,  
 Schmitt,A., Thelings,B., Ritter,E., Ronko,I., Bennett,J., Cardenas  
 M., Gibbons,M., McCann,R., Cole,R., Tsagarashvili,R., Williams,T.,  
 Jackson,Y. and Bowers,Y.

## TITLE

Endocrine Pancreas Consortium

## JOURNAL

Unpublished (2000)

## COMMENT

Other ESTs: Id98f09.y1

## CONTACT

Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138

## TELEPHONE

Tel: 617-495-1812

## FAX

Fax: 617-495-8557

## EMAIL

Email: dmelton@lohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Dr. Hiroshi Inoue  
 (hino@elms.wustl.edu)  
 Seq primer: -40UP from G1bco  
 High quality sequence stop: 399.

## FEATURES

Location/Qualifiers

## SOURCE

1. 599

## ORGANISM

/organism="Homo sapiens"

## DB\_XREF

/db\_xref="taxon:9606"

## CLONE

/clone="IMAGE:5085832"

## CLONE\_LIB

/clone.lib="HR85 Islet"

## TISSUE

/tissue\_type="Purified pancreatic islet"

## LAB\_HOST

/lab\_host="DH10B"

## NOTE

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
 NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
 Size-selected on agarose gel. Average insert size -1kb. 5'  
 XhoI site was destroyed after directional cloning.  
 Amplified once. Contact information: Hiroshi Inoue,  
 Metabolism Div. (Alan Permutt Lab), Washington University  
 School of Medicine, Box 8127, 660 South Euclid Ave., St.  
 Louis, MO 63110, E-mail: hino@elmsgate.wustl.edu, Tel:  
 314-362-1916, Fax: 314-747-2692."

BASE COUNT 133 a 163 c 132 g 171 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 2.99e-102 Length: 599  
 Score: 843.00 Matches: 161  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.41% Indels: 0  
 DB: 13 Gaps: 0

US-10-082-502-20 (1-162) x B1791489 (1-599)

OY 2 ArgserGlnaspLeuHiscysGlyAlaCysArgAlaLeuValaspGluLeuGluTTPG 21  
 |||||||  
 DB 599 AGAGCCAGAGATCTCCACTGTGAGCATGAGGCGCTCTGTGATGACATGAGATGGAA 540  
 |||||||  
 OY 22 IleAlaGlnValaspProlyslsYthrIleGlnMetGlySerPheargIleasProAsp 41  
 |||||||  
 DB 539 ATTGCCAGCTGAGCCCAAGAACACATTCAGATGGGATCTTCCGATCAATCCAGAT 480  
 |||||||



QY 42 GlycerInservValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeu 61  
 Db 479 GGCAGCCAGTGTGTGGAGGTGCTTATGCCCCGCTCAAGGCCACTCACAGACCTG 420  
 QY 62 LeuGluGluIleCysAspArgMetGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 81  
 Db 419 CTGGAGAGATGATGTGACCGGATGAGAGATGAGAGATGAGATGAGATGAGATGAGAT 360  
 QY 82 ArgLysAsnTyrValAlaGluValGlyArgAsnGlyGlyGlyGlyGlyGlyGlyGly 101  
 Db 359 CGGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 300  
 QY 102 GlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlySerIleVal 121  
 Db 299 GGCATCGCATTCACCTCAGATATTAGGCGACCCCTCAAGTTGCGGTGAGACATTTGTG 240  
 QY 122 GluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAsp 141  
 Db 239 GAGGAATACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
 QY 142 LysLeuCysSerTyrArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspGlu 161  
 Db 179 AACTTTGAGTAAGCAGACAGATCTTTGTGACCATGCCCTGCATATCGCATGTAG 120  
 QY 162 Leu 162  
 Db 119 CTA 117  
 RESULT 12  
 LOCUS BQ421827 880 bp mRNA linear EST 23-MAY-2002  
 DEFINITION AGENCOURT\_7801623 NIH\_MGC\_72 Homo sapiens CDNA clone IMAGE:6050568  
 ACCESSION BQ421827  
 VERSION BQ421827.1 GI:21117142  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 880)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: ATCC/DC/PT/DRP  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLMJ3303 row: a column: 01  
 High quality sequence stop: 580.  
 Location/Qualifiers  
 1..880  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6050568"  
 /clone\_lib="NIH\_MGC\_72"  
 /tissue\_type="melanotic melanoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;  
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2 kb. Library constructed by Life  
 Technologies."  
 BASE COUNT 238 a 208 c 256 g 176 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.2e-102 Length: 880  
 Score: 842.00 Matches: 161

Percent Similarity: 99.388  
 Best Local Similarity: 99.388  
 Query Match: 99.29%  
 Db: 14  
 Gaps: 0  
 US-10-082-502-20 (1-162) x BQ421827 (1-880)  
 QY 1 ArgArgSerGluAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 Db 162 CGAGAGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 221  
 QY 21 GluIleAlaGluValAspProLysThrIleGluMetGlySerPheArgIleAspPro 40  
 Db 222 GAAATTTGCCAGAGTGGAGCCCAAGAAAGACATTCAGATGGAGATCTTCCGATCAATCCA 281  
 QY 41 AspGlySerGluSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 Db 282 GATGGCAGACCATGAGTGGTGGAGGTGCTTATGCGGCTCAGAGGCCCACTTCACAG 341  
 QY 61 LeuLeuGluGluIleCysAspArgMetGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 80  
 Db 342 CTGCTGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 401  
 QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyGlyGlyGlyGlyGlyGlyGly 100  
 Db 402 CATGCCAAGAACTACGTACGTGTAGTGGCCGTATGAGAACTCAGTGAACCTGACCTTA 461  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlySerIle 120  
 Db 462 CAAGCAGCCCAATCGATCAGATATTAGCGGACCCCTCAAGTTGCGGTGAGAGACATTT 521  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 Db 522 GTGGAGGAATACGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 581  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 Db 582 GACAAACTTTGAGTAAGCAGACAGATCTTTGTGACCATGCCCTGCATATCGCATGAT 641  
 QY 161 GluLeu 162  
 Db 642 GAGCTA 647  
 RESULT 13  
 LOCUS B1550436 783 bp mRNA linear EST 05-SEP-2001  
 DEFINITION 603192905F1 NIH\_MGC\_95 Homo sapiens CDNA clone IMAGE:5264328 5',  
 ACCESSION B1550436  
 VERSION B1550436  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 783)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki  
 Toshitaki and Piero Carninci (RIKEN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLMJ1666 row: 1 column: 01  
 High quality sequence stop: 730.  
 Location/Qualifiers  
 1..783

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:5264328"  
/clone_lib="NIH_MGC_95"  
/tissue_type="hippocampus"  
/lab_host="DH10B"  
/note="Organ: brain; Vector: pBluescript (modified  
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag  
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.5 kb and  
normalized to 80% 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIMH/NHGRI, National  
Institutes of Health). Note: this is a NIH_MGC Library."  
BASE COUNT      209 a      189 c      225 g      160 t  
ORIGIN  
  
Alignment Scores:  
Pred. No.:      1,12e-101      Length:      783  
Score:      840.00      Matches:      161  
Percent Similarity:      99.38%      Conservative:      0  
Best Local Similarity:      99.38%      Mismatches:      1  
Query Match:      99.06%      Indels:      0  
DB:      13      Gaps:      0  
  
US-10-082-502-20 (1-162) x B1550436 (1-783)  
OY      1      ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
      |||||  
Db      144      CGGAGAGCCAGGATCTCCACGTGTGAGCATGACGGCTCTGTGTGATGACATGATG 203  
OY      21      GluIleAlaGlnValAspProLysLysThrIleGlnMetClySerPheArgIleAsnPro 40  
      |||||  
Db      204      GAAATTGCCAGGTGAGCCCAAGAACCATTCAGATGGATCTTCGGATCAATCA 263  
OY      41      AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
      |||||  
Db      264      GATGGACCCAGTCAGTGTGAGGTGCTTATGCCCTTAAAGGCCCACTCACAGAG 323  
OY      61      LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 80  
      |||||  
Db      324      CTGCTGGAGAGATATGTACCGGATGAGCATGATGGGACATGATCTTCACAC 383  
OY      81      HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
      |||||  
Db      384      CATCGCAAACTACTACTGATGTAGTGGCCGGAATGAGAAATCCAGTGAATGACCTA 443  
OY      101      GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
      |||||  
Db      444      CAAGGATCCGAAATGACATGATATATAGCGGACACCTCAAGTTTCCGTGTGAGACAT 503  
OY      121      ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
      |||||  
Db      504      GTGGAGGATACGAGATTAATCAATCTTATTTCCCGAGAGGCTGACATGTTTAA 563  
OY      141      AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
      |||||  
Db      564      GACAAACTTTGAGATGAGGAGACAGATCTTTGTGACCATGACCTGACATATCGATGAT 623  
OY      161      GluLeu 162  
      |||||  
Db      624      GAGCTA 629  
  
RESULT 14  
BO421069      868 bp      mRNA      linear      EST 23-MAY-2002  
LOCUS      AGENCOURT.7911956 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:5009949  
DEFINITION      5', mRNA sequence.  
ACCESSION      BO421069  
VERSION      BO421069.1 GI:21116384  
KEYWORDS      EST.  
SOURCE      human.
```

```
ORGANISM      Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE      1 (bases 1 to 868)  
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.  
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL      Unpublished (1999)  
COMMENT      Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-riemail.nih.gov  
Tissue Procurement: DCTD/DRP/Gazdar  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM13197 row: d column: 14  
High quality sequence stop: 542.  
Location/Qualifiers  
1..868  
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:6009949"  
/clone_lib="NIH_MGC_68"  
/tissue_type="large cell carcinoma"  
/lab_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;  
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.8 kb. Library constructed by Life  
Technologies."  
BASE COUNT      232 a      203 c      257 g      168 t      8 others  
ORIGIN  
  
Alignment Scores:  
Pred. No.:      6.09e-101      Length:      868  
Score:      835.00      Matches:      160  
Percent Similarity:      98.77%      Conservative:      0  
Best Local Similarity:      98.77%      Mismatches:      2  
Query Match:      98.47%      Indels:      0  
DB:      14      Gaps:      0  
  
US-10-082-502-20 (1-162) x BO421069 (1-868)  
OY      1      ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
      |||||  
Db      190      CGGAGAGCCAGGATCTCCACGTGTGAGCATGACGGCTCTGTGTGATGACATGATG 249  
OY      21      GluIleAlaGlnValAspProLysLysThrIleGlnMetClySerPheArgIleAsnPro 40  
      |||||  
Db      250      GAAATTGCCAGGTGAGCCCAAGAACCATTCAGATGGATCTTCGGATCAATCA 309  
OY      41      AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
      |||||  
Db      310      GATGGACCCAGTCAGTGTGAGGTGCTTATGCCCTTAAAGGCCCACTCACAGAG 369  
OY      61      LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAsnProSerThr 80  
      |||||  
Db      370      CTGCTGGAGAGATATGTACCGGATGAGCATGATGGGACATGATCTTCACAC 429  
OY      81      HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
      |||||  
Db      430      CATCGCAAACTACTACTGATGTAGTGGCCGGAATGAGAAATCCAGTGAATGACCTA 489  
OY      101      GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
      |||||  
Db      490      CAAGGATCCGAAATGACATGATATATAGCGGACACCTCAAGTTTCCGTGTGAGACAT 549  
OY      121      ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
      |||||  
Db      550      GTGGAGGATACGAGATTAATCAATCTTATTTCCCGAGAGGCTGACATATCGATGAT 609  
OY      141      AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
      |||||
```

Db 610 GACAACTTTGACGTACGACAGACATCTTTGTGACCATGCCCTGCACATATGCATGAT 669  
 QY 161 GIUleu 162  
 Db 670 GAGCTA 675

RESULT 15  
 BG704443

LOCUS  
 DEFINITION BG704443 791 bp mRNA linear EST 07-MAY-2001  
 602686633F1 NIH\_MGC\_95 Homo sapiens CDNA clone IMAGE:4821008 5',  
 mRNA sequence.  
 ACCESSION BG704443  
 VERSION BG704443.1 GI:13977790  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@rs-remail.nih.gov  
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM10727 row: a column: 09  
 High quality sequence stop: 785.

FEATURES  
 Source  
 Location/Qualifiers  
 1..791

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4821008"  
 /clone\_1lb="NIH\_MGC\_95"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH108"  
 /note="Organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtagag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
 size-selected for average insert size 2.5 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NHGRI/NHGRJ, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 187 a 196 c 256 g 152 t

# Alignment Scores:

Pred. No.: 9.83e-101 Length: 791  
 Score: 833.00 Matches: 160  
 Percent Similarity: 99.38% Conservative: 1  
 Best Local Similarity: 98.77% Mismatches: 1  
 Query Match: 98.23% Indels: 0  
 DB: 12 Gaps: 0

US-10-082-502-20 (1-162) x BG704443 (1-791)

QY 1 ArgATgSerGlnAspLeuNHicysGlyAlaCysArgAlaLeuValAspGluLeuGluTrr 20  
 Db 259 CGGAGAGACCCAGGATCTCCAGATGTCGACAGCGCTCTGGTGATGAAGTGA 318  
 QY 21 GluLeaGlnValAspProGlySerThrIleGlnMetGlySerPheArgIleasnPro 40  
 Db 319 GAATGTGCCAGGTGGACCCCAAGAGACCATTCAGATGGGATCTTCGGATCAATCCA 378  
 QY 41 AspGlySerGlnSerValIcIuValProTyrAlaArgSerGluAlaHisLeuThrGlu 60

Db 379 GATGGCAGCCAGTCAGTGGTGGAGGTGCCCTTATGCCCTCAGAGGCCACACCTCACAGAG 438  
 QY 61 LeuLeuGluGluIleCysAspArgMetIysGluTyrGlyGluGlnIleAspProSerThr 80  
 Db 439 CTGCTGGAGAGATATGTGACCGGATGMAAGAGATATGGGAAACAGATTGATCTTCACACC 498  
 QY 81 HisArgIysAsnTyrValArgValValGlyArgasnGlyIleSerSerGluLeuAspLeu 100  
 Db 499 CATGCAAGAACTACGTACGTAGTGGCCGGAATGAGAAATCCAGTCACTGAGACCTA 558  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 Db 559 CAAAGCATCCAAATCCACTCAGTATTTACGGCACCCCTCAAGTTGGCTGTGAGACAT 618  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValIys 140  
 Db 619 GTGGAGGAATAGAGAGATGAATCATTTGAATTTCCCGAGAGCTGCACAAATTTTAA 678  
 QY 141 AspLysLeuCysSerIysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 Db 679 GACAACTTTGAGATAGTAAAGCAACAGATCTTTGTGACCATGCCCTGCACATATCGCATGAT 738  
 QY 161 GIUleu 162  
 Db 739 GAGCTA 744

Search completed: March 31, 2003, 04:51:11  
 Job time: 1041.65 secs





Alignment Scores:  
Pred. No.: 2,43e-102 Length: 814  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-082-502-20 (1-162) x US-09-864-864-309 (1-814)

OY 1 ARGARGSGLNAPLEUHSGLYALACYSARGALALEUVALASPLUENGLUTP 20  
DB 205 CGAGAGAGCCAGGATCTCCACTGTGGAGCATGGGCTTGTTGGATGAGATG 264  
OY 21 GLLLEALAGLVALAPPROLYSLYTHRIEGLMCTGLYSEPHEARGLAENPRO 40  
DB 265 GAAATGGCCAGGTGAGCCCAAGAACACATTCAGATGGATCTTCCGATCAATCCA 324  
OY 41 ASBGLYSERGLNSERVALVALGLUVALPROTYRALARGSERGLUALHISLEUTHRLU 60  
DB 325 GATGGCAGCCAGCATGTTGGAGTGCCCTTATGCCCCCTCAGAGGCCACCTCACAGAG 384  
OY 61 LEULEUGLGLULLECYASAPARGMETLYSGLUYRGLYGLULINLEASPROSERTHR 80  
DB 385 CTGCTGAGAGCATATGTGACCGGATGAGGAGTATGGGAAACAGATTGCTTCCACC 444  
OY 81 HLSARGLYASANTYRVALARGVALVALGLYARGANGLYGLUSERSERGLUEASPLEU 100  
DB 445 CATCGCAAGACATACGTACGTAGTGGGCCGGAATGGAGAAATCCAGTAACTGGACCTA 504  
OY 101 GINGLYTLEARGTLEASPSERASPILESERGLYTHLEULYSPHEALACYSGLUSERLIE 120  
DB 505 CAAGGATCCGATCCAGTACCATGATTACGGCACCCCTCAAGTTCTGCTGAGACCAT 564  
OY 121 VALGLULTYRGLUASPLUENLEGLUPHESERARGSLUALHASPANVALLYS 140  
DB 565 GTGGAGAGATACAGACATCACTGAAATCTTTTCCGAGAGGCTGACATGTTAAA 624  
OY 141 AEPPLYSLEUCYSERLYARGLTHRAPLEUCYSAPHLAHLAHLISLESERHASP 160  
DB 625 GACAACTTGTGAGTAAGCAACAGATCTTTGTGACCATGCCCCCTGCACATATGCATGAT 684  
OY 161 GLULEU 162  
DB 685 GAGCTA 690

RESULT 2  
US-10-028-072-173  
; Sequence 173, Application US/10028072  
; Publication No. US20030004311A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang  
; TITLE OF INVENTION:  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/028,072  
; CURRENT FILING DATE: 2001-12-19

;; PRIOR APPLICATION NUMBER: 60/049911  
;; PRIOR FILING DATE: 1997-06-18  
;; PRIOR APPLICATION NUMBER: 60/056974  
;; PRIOR FILING DATE: 1997-08-26  
;; PRIOR APPLICATION NUMBER: 60/059113  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/059115  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/059117  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/059122  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/059184  
;; PRIOR FILING DATE: 1997-09-17  
;; PRIOR APPLICATION NUMBER: 60/059263  
;; PRIOR FILING DATE: 1997-09-18  
;; PRIOR APPLICATION NUMBER: 60/059352  
;; PRIOR FILING DATE: 1997-09-19  
;; PRIOR APPLICATION NUMBER: 60/059588  
;; PRIOR FILING DATE: 1997-09-19  
;; PRIOR APPLICATION NUMBER: 60/059836  
;; PRIOR FILING DATE: 1997-09-24  
;; PRIOR APPLICATION NUMBER: 60/062250  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/062285  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/062287  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/062814  
;; PRIOR FILING DATE: 1997-10-24  
;; PRIOR APPLICATION NUMBER: 60/062816  
;; PRIOR FILING DATE: 1997-10-24  
;; PRIOR APPLICATION NUMBER: 60/063045  
;; PRIOR FILING DATE: 1997-10-24  
;; PRIOR APPLICATION NUMBER: 60/063082  
;; PRIOR FILING DATE: 1997-10-31  
;; PRIOR APPLICATION NUMBER: 60/063127  
;; PRIOR FILING DATE: 1997-10-24  
;; PRIOR APPLICATION NUMBER: 60/063327  
;; PRIOR FILING DATE: 1997-10-27  
;; PRIOR APPLICATION NUMBER: 60/063329  
;; PRIOR FILING DATE: 1997-10-27  
;; PRIOR APPLICATION NUMBER: 60/063550  
;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/063561  
;; PRIOR FILING DATE: 1997-10-28  
;; PRIOR APPLICATION NUMBER: 60/063704  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063733  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063735  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063738  
;; PRIOR FILING DATE: 1997-10-29  
;; PRIOR APPLICATION NUMBER: 60/063755  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064248  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/064809  
;; PRIOR FILING DATE: 1997-11-07  
;; PRIOR APPLICATION NUMBER: 60/065186  
;; PRIOR FILING DATE: 1997-11-12  
;; PRIOR APPLICATION NUMBER: 60/065846  
;; PRIOR FILING DATE: 1997-11-17  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/066453  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066511  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/066770  
;; PRIOR FILING DATE: 1997-11-24  
;; PRIOR APPLICATION NUMBER: 60/069212

PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083222  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086414  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086430  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088730  
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088741  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

## Alignment Scores:

Pred. No.: 4,2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-082-502-20 (1-162) x US-10-028-072-173 (1-1210)

QY 1 ARGATSERGINASPLENHISCYGLYALACYSATGALALEUVALASPLULEUGLUTP 20  
DB 603 CGGAGGACCGAGATCTCCACGTGTGAGCATGCCGTCTCGTGAGTACGATGCG 662  
QY 21 GLULEALAGLVALASPROLYSLYSTRILIEGIMETGLYSEPHARGLEASPRO 40  
DB 663 GAAATTTGCCAGGTGGACCCCAAGAGACCATTCAGATGGGATCTTCCGGATCAATCA 722  
QY 41 ASPLYSERGINSERVALALGULVALPROTYRALAARGSERGLUALAHISLEUTHGLU 60  
DB 723 GATGACAGCAGATCAGTGTGTGAGGTGCTTATGCCCGCTCGAGGCCACATCAGAG 782  
QY 61 LEULEUGLUGLULIECYASAPRAGMETLYSGULTYGTYGLUGLILIAAPRROSETHR 80  
DB 783 CTGCTGGAGGAGATATGTCACCGGATGAAGGAGTGGGGAACAGATTGATCTCCACC 842  
QY 81 HISARGLYASANTYVALARGVALALGARGANGLYGLUSERGGLULEUASPLEU 100  
DB 843 CATCCAGAACCTTACGTACGTGTACTGGGCCGAATGAGAACTTCCAGTGAACCTGA 902  
QY 101 GINGLYILEARGLIEASPSERASPLIESERGLYTHLEULYRPHALCYSLUSEITIE 120  
DB 903 CAAGCATCCGAATGAGTACATGATTTAGCGGCACCCCAAGTTGGCGGTGAGACATT 962  
QY 121 VALGLIGULTYRGLUASPLULEULIEGLUPHEPSESERATGGLUALASPSANVALYS 140  
DB 963 GTGGAGGAATAGAGAGATGAACCTCATTTATCTTTCCCGAGAGCTGCACAACTGTA 1022  
QY 141 ASPLYSLEUCYSERLYARGTHASPLEUCYASAPHIALALEUHLIALESERHASP 160  
DB 1023 GCAAACTTTGCGATGAGGAGAACAGATCTTTGTGCCATGCCCTGCACATATGCGATAT 1082



OY 161 Glutelu 162  
 DB 1083 GAGCTA 1088

## RESULT 3

US-10-121-049-173  
 ; Sequence 173, Application US/10121049  
 ; Publication No. US20030022239A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Collin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3330R1C17  
 ; CURRENT APPLICATION NUMBER: US/10/121,049  
 ; CURRENT FILING DATE: 2002-04-12  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 173  
 ; LENGTH: 1210  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-121-049-173

## Alignment Scores:

Pred. No.: 4,2e-102 Length: 1210  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-121-049-173 (1-1210)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 DB 603 CGAGGAGCCAGGATCTCCACTCTGCGCATGCGAGGCTCTGGTGAACAGAAATGG 662  
 OY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 DB 663 GAATATGCCAGATGTGACCCCAAGAACCATTCAGATGGGATCTTCCGGAATCAATCCA 722  
 OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 DB 723 GATGGAGCCAGATCAAGTGGAGGTGCTTATGCGCGCTCAAGAGGCCCTCACAGAG 782  
 OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 783 CTCTGTGAGAGATATGTGACCGAGTAAGAGATATGGGAGAACAGATGATCTTCACAC 842  
 OY 81 HisArgLysAsnTyrValAlaArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100  
 DB 843 CATCGCAAGACTACAGCTAGTGGGCGGAAATGGAAATCCAGTGAAGTGAAGCTTA 902  
 OY 101 GlnGlyTLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 903 CAAGGATCGAATCGATCGATATTAAGGAGGACCCCTCAAGTTTGCTGTGAGAGCAT 962

OY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGlnAlaAspAsnValLys 140  
 DB 963 GTGAGGAATACAGAGATGATCAATTCATTTCCCGAGAGGCTGACAAATGTTAAA 1022  
 OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 1023 GACAAACTTTGCACTAAGCAGACAGATCTTGTGACCATGCTCCCTGCACATATGCGATGAT 1082  
 OY 161 Glutelu 162  
 DB 1083 GAGCTA 1088

## RESULT 4

US-10-123-904-173  
 ; Sequence 173, Application US/10123904  
 ; Publication No. US20030022338A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Collin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3330R1C54  
 ; CURRENT APPLICATION NUMBER: US/10/123,904  
 ; CURRENT FILING DATE: 2002-04-16  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 173  
 ; LENGTH: 1210  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-123-904-173

## Alignment Scores:

Pred. No.: 4,2e-102 Length: 1210  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-123-904-173 (1-1210)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 DB 603 CGAGGAGCCAGGATCTCCACTGTGAGAGATGCAAGGCTTGTGTGATGAATGAATGG 662  
 OY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 DB 663 GAATATGCCAGATGTGACCCCAAGAACCATTCAGATGGGATCTTCCGGAATCAATCCA 722  
 OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60  
 DB 723 GATGGAGCCAGATCAAGTGGAGGTGCTTATGCGCGCTCAAGAGGCCCTCACAGAG 782  
 OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 783 CTCTGTGAGAGATATGTGACCGAGTAAGAGATATGGGAGAACAGATGATCTTCACAC 842

QY	81	HISRTGYASNNYTVaLaRGVaLVaIGLYARGASNGLYSERSerGLUeuSpleu	100
Db	843	CATGCGAAGAACTACGACTGTGTTGGGCGGGAATGGGAATCCAGTAACTGGACCTA	902
QY	101	GLNGLYIleatgIleaspsSerapIleSerGlyThrLeuLysPheALaCysGLuSerIle	120
Db	903	CAAGCGCTCCGAATCGCACTCGATATATTAGCGGCAACCCCTCAAGTTTGGCTGAGAGCATTT	962
QY	121	ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluValAspArgValLys	140
Db	963	GTTGGAGGAAATACGAGATGCACTCATTTGAATTTCTTTCCGAGAGGCTATACATATGTAA	1022
QY	141	AspLysLeuCysSerLysArgThrAspLeuCysAspPheALaLeuIleIleSerHisasp	160
Db	1023	GACAAATTTGCAGTAAAGCAGACAGATCTTTGTGACCAATGCCCTGCACATATCGCATGT	1082
QY	161	GluLeu 162	
Db	1083	GAGCTA 1088	

Seq ID	Seq	Length
1	ATGATGSEGGTGAAPRLNLSHCYSGLYALACYSARGALALEUVALASPRILUENGLUTP	20
603	CGAGAGACCCAGGATCTCCACTGTCGAGACATGCAAGGGCTCTGGTGGATGAACTACAAAGG	662
21	GLUILEAGLVALAASPRLYLSYTHRIEGLIMETGLYSERPHEARGLLEASPRO	40
663	GAATATGCCCGAGTGGGACCCCAAGAGACCATTGACATGGGATCTTCCGGATCAATCCA	722
41	ASPGLSERGLNSERVALVALGLUVALPROTYNALARGSERGLNALHISLEUTHRGU	60
723	GATGGCACCCAGTGAAGTGGGAGAGTGCCATTATGCCCCGCTCAAGAGGCCACCTCACAGAG	782
61	LEULEUGLUGLUILECYASAPARGMETLYSGLUTYRGLYGLUINLLEASPROSERTHR	80
783	CTGCTGGAGAGGATATGTGTGACCCGATAGAGAGATATGGGACACAGTTATCTCTCCACC	842
81	HLSARGLYASAENYTRVALARGVALGLYARGASNGLYGLUSERSERGLUENASPLEU	100
843	CATGCGAAGACACTACGTAACGTATGATGGCGCGAATGGAATCCAGTGACTGGACCTA	902
101	GLINGLYLEARGLLEASPRSERASPILLESERGLYTHLEULYSPHEALAEYSGLUSERITL	120
903	CNAGGCATCCCAATCGACTAGATATTAGCGGCCCTCAAGTTGCGTGTGAGAGCATY	962
121	VALGLUGLUTYRGLUASPRGLUENILEGLIUPHESERFARGLUNLASPASVALYLS	140
963	GTGAGAGAAATCGAGAGATGAACATCATTTGATTTTCCCGAAGAGCTGACAAATGTTAA	1022
141	ASPLYSLEUCYSERYASRGTRHASPLEUCYASAPHLALALEUHLISLESERHASP	160
1023	GACAAACTTTCAGTAAAGCGAACAGATCTTGTGACCATGCCCTGCACATATTCGATGAT	1082
161	GLULEU 162	
1083	GAGCTA 1088	

RESULT 7

US-10-176-918-173

Sequence 173, Application US/10176918

Publication No. US2003027275A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Matanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3330R1C382

CURRENT APPLICATION NUMBER: US/10/176,918

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 173

LENGTH: 1210

TYPE: DNA

ORGANISM: Homo Sapien

US-10-176-918-173

Alignment Scores:

Pred. No.: 4.2e-102

Length: 1210

Percent Similarity:	848.00	Matches:	162
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-082-502-20 (1-162) x US-10-176-918-173 (1-1210)

QY	1	ArgAgsSerGlnAspPleuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp	20
DB	603	CGAGAGACCACCGAGATCTCCACTGTGGAGCAGCAGGGCTGTGGGTGTAACATGATG	662
QY	21	GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro	40
DB	663	GAAATTGCCCGAGGTGGAGCCCAAGAGACCTTACATGGGATCTTCCGATCAATCCA	722
QY	41	AspGlySerGlnSerValAlaGluValProTrpAlaArgSerGluAlaHisLeuArgIle	60
DB	723	GATGGCACCAGCATGCTGTGTGAGAGTGGCTTATGCCCGTCAGAGGCCACACAGAG	782
QY	61	LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr	80
DB	783	GTGCTGGAGGAGATATGTGACCGGATGAMGAGATGTGGGAAACGATGTATCTCCACC	842
QY	81	HisArgLysAsnTrpValArgValAlaGlyArgAsnGlyGlnSerSerGluLeuAspLeu	100
DB	843	CATCCCAAGAACTACGTCACGTGTGTGTGTGGCCGGAATGAGATTCATGAACTGGACCTA	902
QY	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
DB	903	CAAGGCATCCGAAATCGACTCAGATATTAGCGGCACCTCAAGTTTGGGTGTGAGACAT	962
QY	121	ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys	140
DB	963	GTGGAGGAATCGAGAGAACTCATGAAATCTTTTCCGAGAGGCTGACAAATGTTAAA	1022
QY	141	AspLysLeuCysSerLysArgTrpHisAspLeuCysAspHisAlaLeuHisIleSerHisAsp	160
DB	1023	GACAAACTTGACGTACCGAAGACAGATCTTTGTACCATGCCCTGCACATATCGCATGAT	108
QY	161	GluLeu 162	
DB	1083	GAGCTA 1088	

RESULT 8  
US-10-176-921-173  
: Sequence 173, Application US/10176921  
: Publication No. US20030027276A1  
: GENERAL INFORMATION:  
: APPLICANT: Baker, Kevin P.  
: APPLICANT: Beresini, Maureen  
: APPLICANT: DeForge, Laura  
: APPLICANT: Desnoyers, Luc  
: APPLICANT: Filvaroff, Ellen  
: APPLICANT: Gao, Mei-Qiang  
: APPLICANT: Gerritsen, Mary E.  
: APPLICANT: Goddard, Audrey  
: APPLICANT: Godowski, Paul J.  
: APPLICANT: Gurney, Austin L.  
: APPLICANT: Sherwood, Steven  
: APPLICANT: Smith, Victoria  
: APPLICANT: Stewart, Timothy A.  
: APPLICANT: Tunas, Daniel  
: APPLICANT: Watanabe, Colin K  
: APPLICANT: Wood, William  
: APPLICANT: Zhang, Zemin  
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
: TITLE OF INVENTION: ACIDS ENCODING THE SAME  
: FILE REFERENCE: P3330R1C28  
: CURRENT APPLICATION NUMBER: US/10/176,921  
: CURRENT FILING DATE: 2002-06-20  
: Prior Application removed - See File Wrapper or Palm  
: NUMBER OF SEQ ID NOS: 550

SEQ ID NO 173  
 LENGTH: 1210  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-176-921-173

## Alignment Scores:

Pred. No.:	4,2e-102	Length:	1210
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-082-502-20 (1-162) x US-10-176-921-173 (1-1210)

```

OY 1 ArgArGserGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
    |||||
DB 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGGATGAACTAGAAATGG 662
OY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||
DB 663 GAAATTCGCCAGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGATCAATCCA 722
OY 41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||||
DB 723 GATGGCAGCCAGTCAGTGGTGGAGTGCTCTTATGCCGCTCAGAGGCCACCTCAGAGAG 782
OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||||
DB 783 CTCTGTGAGAGATATGTGACCGGATGAAGAGTATGGGAAACAGTTGATCTTCCACC 842
OY 81 HisArgLysAsnTyrValArgValIleGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
    |||||
DB 843 CATCGCAAGAACTACGTAGCTGTAGTGGCCCGGAATGGAATCCCACTGAGACTGA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||
DB 903 CAAGGCATCCGAAATCGACATATTAACGGCACCCCTCAAGTTTCCGTGTAGACCAT 962
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||||
DB 963 GTGGAGGAATACGAGGATGAATCATTAATCTTTCCCGAGAGGCTGACAAATGTTAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
    |||||
DB 1023 GACAAACTTTCAGTAGAGCAACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 1082
OY 161 GluLeu 162
    |||||
DB 1083 GAGCTA 1088
  
```

RESULT 9  
 US-10-137-865-173  
 Sequence 173, Application US/10137865  
 Publication No. US20030032155A1  
 GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: DeForge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William

APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 FILE OF INVENTION: ACIDS ENCODING THE SAME  
 FILE REFERENCE: P3330R1C154  
 CURRENT APPLICATION NUMBER: US/10/137, 865  
 PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
 NUMBER OF SEQ ID NOS: 550  
 SEQ ID NO 173  
 LENGTH: 1210  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-137-865-173

## Alignment Scores:

Pred. No.:	4,2e-102	Length:	1210
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-082-502-20 (1-162) x US-10-137-865-173 (1-1210)

```

OY 1 ArgArGserGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
    |||||
DB 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGAGGCTGTGGTGGATGAACTAGAAATGG 662
OY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||
DB 663 GAAATTCGCCAGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGATCAATCCA 722
OY 41 AspGlySerGlnSerValIleGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||||
DB 723 GATGGCAGCCAGTCAGTGGTGGAGTGCTCTTATGCCGCTCAGAGGCCACCTCAGAGAG 782
OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||||
DB 783 CTCTGTGAGAGATATGTGACCGGATGAAGAGTATGGGAAACAGTTGATCTTCCACC 842
OY 81 HisArgLysAsnTyrValArgValIleGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
    |||||
DB 843 CATCGCAAGAACTACGTAGCTGTAGTGGCCCGGAATGGAATCCCACTGAGACTGA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||
DB 903 CAAGGCATCCGAAATCGACATATTAACGGCACCCCTCAAGTTTCCGTGTAGACCAT 962
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||||
DB 963 GTGGAGGAATACGAGGATGAATCATTAATCTTTCCCGAGAGGCTGACAAATGTTAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
    |||||
DB 1023 GACAAACTTTCAGTAGAGCAACAGATCTTTGTGACCATGCCCTGCACATATGCATGAT 1082
OY 161 GluLeu 162
    |||||
DB 1083 GAGCTA 1088
  
```

RESULT 10  
 US-10-140-474-173

Sequence 173, Application US/10140474  
 Publication No. US20030032156A1  
 GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Beresini, Maureen  
 APPLICANT: DeForge, Laura  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Filvaroff, Ellen  
 APPLICANT: Gao, Wei-Qiang  
 APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C162
; CURRENT APPLICATION NUMBER: US/10/140,474
; PRIOR APPLICATION REMOVED - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-173

Alignment Scores:
Pred. No.: 4.2e-102      Length: 1210
Score: 848.00           Matches: 162
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 9                         Gaps: 0

US-10-082-502-20 (1-162) x US-10-140-474-173 (1-1210)

QY 1 ArgArgSerGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGGAGGAGCCAGATCTCCACTGTGAGCATGCGAGGCTGTGGTGAATGAGAAATG 662
QY 21 GlnIleAlaGlnValAspProLysLysThrIleGlnMetLysSerPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCCGGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60
DB 723 GATGGACGCCAGTCACTGCTGAGAGTCCCTTATGCCCGCTCAAGGCCCTCAGAGAG 782
QY 61 LeuLeuGlnGluIleCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80
DB 783 CTCTGAGAGAGATATGTGACCGGATGAAGAGATAGGGGAGACATGATCTTCACAC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTAGCTAGCTGATGGGCGGAAATGGAGAAATCCAGTGAATCGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAGGCGATCGAATCGACTAGATATTAGGGGACCCCTCAAGTTTGGCTGAGAGCATT 962
QY 121 ValGluGluTyrGlnAspGluLeuIleGluPhePheSerArgGlnAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGAGATGAATCATTTGTTCCCGAGAGAGGTGACATGTTTAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
DB 1023 GACAAACTTTGACGATGAAGCAACAGATCTTTGTGACCAATGCCCTGCAATATCGCATGAT 1082
QY 161 GluLeu 162
DB 1083 GACCTA 1088

RESULT 11
US-10-142-431-173
; Sequence 173, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-173

Alignment Scores:
Pred. No.: 4.2e-102      Length: 1210
Score: 848.00           Matches: 162
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 9                         Gaps: 0

US-10-082-502-20 (1-162) x US-10-142-431-173 (1-1210)

QY 1 ArgArgSerGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGGAGGAGCCAGATCTCCACTGTGAGCATGCGAGGCTGTGGTGAATGAGAAATG 662
QY 21 GlnIleAlaGlnValAspProLysLysThrIleGlnMetLysSerPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCCGGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60
DB 723 GATGGACGCCAGTCACTGCTGAGAGTCCCTTATGCCCGCTCAAGGCCCTCAGAGAG 782
QY 61 LeuLeuGlnGluIleCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80
DB 783 CTCTGAGAGAGATATGTGACCGGATGAAGAGATAGGGGAGACATGATCTTCACAC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTAGCTAGCTGATGGGCGGAAATGGAGAAATCCAGTGAATCGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAGGCGATCGAATCGACTAGATATTAGGGGACCCCTCAAGTTTGGCTGAGAGCATT 962
QY 121 ValGluGluTyrGlnAspGluLeuIleGluPhePheSerArgGlnAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGAGATGAATCATTTTTCCTCCGAGAGGTGACATGTTTAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
DB 1023 GACAAACTTTGACGATGAAGCAACAGATCTTTGTGACCAATGCCCTGCAATATCGCATGAT 1082
QY 161 GluLeu 162

```

Db 1083 GAGCTA 1088  
|||||  
RESULT 12  
US-10-143-114-173  
; Sequence 173, Application US/10143114  
; Publication No. US20030036180A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C211  
; CURRENT APPLICATION NUMBER: US/10/143,114  
; PRIOR FILING DATE: 2002-05-09  
; PRIOR APPLICATION removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-143-114-173  
  
Alignment Scores:  
Pred. No.: 4.2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 9  
US-10-082-502-20 (1-162) x US-10-143-114-173 (1-1210)  
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
Db 603 CGAGAGAGCCAGAGATCTCCAGTGTGAGCATGCAAGGCGCTGTGTGATGAACTAGAAATGG 662  
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
Db 663 GAAATTGCCAGGTGAGATGCCCAAGAGACCAATTCAGATGGATCTTCCGATCAATCA 722  
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
Db 723 GATGGACGCCAGTCAAGTGTGAGAGTTCCTTATGCCGCTCAGAGGCCCACTCACAAGAG 782  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
Db 783 CTGCTGGAGAGATATGTGACCGAGATGAGAGATGGGAAACAGATTGATCTCCACAC 842  
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100  
Db 843 CATCCCAAGAACTACGTACGTACTACTGGCCGGAATGGAGAAATCCAGTGAACCTGACCTA 902  
QY 101 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
Db 903 CAAGGCATCCGATGATGATATATAGCGACACCTCAAGTTTGGCTGTGAGAGCAT 962  
QY 121 ValGluGluTyrGluAspGluLeuLeuLeuPheSerArgGluAlaAspAsnValLys 140

Db 963 GTGAGGAATACGAGATGACATCATTTCTTTCCGAGAGCGCTGACAAATGTTAA 1022  
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
Db 1023 GACAACCTTTGCACTAAGCAGAACAGACATCTTGTGACCATGCGCTGACATATCGCATGAT 1082  
QY 161 GluLeu 162  
Db 1083 GAGCTA 1088  
|||||  
RESULT 13  
US-10-140-002-173  
; Sequence 173, Application US/10140002  
; Publication No. US20030037623A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C39  
; CURRENT APPLICATION NUMBER: US/10/140,002  
; PRIOR FILING DATE: 2002-05-06  
; PRIOR APPLICATION removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-002-173  
  
Alignment Scores:  
Pred. No.: 4.2e-102 Length: 1210  
Score: 848.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-10-082-502-20 (1-162) x US-10-140-002-173 (1-1210)  
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
Db 603 CGAGAGAGCCAGAGATCTCCAGTGTGAGCATGCAAGGCGCTGTGTGATGAACTAGAAATGG 662  
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
Db 663 GAAATTGCCAGGTGAGATGCCCAAGAGACCAATTCAGATGGATCTTCCGATCAATCA 722  
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
Db 723 GATGGACGCCAGTCAAGTGTGAGAGTTCCTTATGCCGCTCAGAGGCCCACTCACAAGAG 782  
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
Db 783 CTGCTGGAGAGATATGTGACCGAGATGAGAGATGGGAAACAGATTGATCTCCACAC 842  
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100

```
|||||
Db 843 CATGCGAAGACTGACGTGATGAGCGCGAATGAGAAATCCATGCACTGAGACTTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 903 CAGGCACTTCGAGATCGATCGATGATATTAGCGGACCCCTCAAGTTGGGTGTGAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
Db 963 GTGGAGGAATACGAGATGCAATCATTCATTCCTTTCCGAGAGCGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 1023 GACAACTTTCAGATTAAGCAACAGATCTTTGTGACCATCCCTGCACATATGCAATGAT 1082
QY 161 GluLeu 162
Db 1083 GAGCTA 1088

RESULT 14
US-10-142-419-173
; Sequence 173, Application US/10142419
; Publication No.: US20030044945A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Zhang, Zemin
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C244
; CURRENT APPLICATION NUMBER: US/10/142,419
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-419-173

Alignment Scores:
Pred. No.: 4,2e-102 Length: 1210
Score: 848.00 Matches: 162
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-142-419-173 (1-1210)
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20
Db 603 CGAGAGAGCCAGATCTCCCTGTCGAGATGCAAGGCTCTGTTGATGATGAATGAGATGG 662
QY 21 GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 663 GAAATTCGCCAGGTGAGCCCAAGAGACCATTCAGATGGGATCTTTCCGAGCAATCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
```

```
|||||
Db 723 GATGCGAGCCAGATCAGTGAAGGAGGTCCCTTATGCCCCGTCAGAGGCCACCTCACAGAG 782
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 783 CTGCTGAGAGAGATATATGTGACCGGATAGAGATGATGGGAAACAGATTTCTTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100
Db 843 CATGCGAAGAACTACGATGATGATGAGCGCGAATGAGAAATCCATGCACTGAGACTTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 903 CAGGCACTTCGAGATCGATCGATGATATTAGCGGACCCCTCAAGTTGGGTGTGAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
Db 963 GTGGAGGAATACGAGATGCAATCATTCATTCCTTTCCGAGAGCGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 1023 GACAACTTTCAGATTAAGCAACAGATCTTTGTGACCATCCCTGCACATATGCAATGAT 1082
QY 161 GluLeu 162
Db 1083 GAGCTA 1088

RESULT 15
US-10-123-262-173
; Sequence 173, Application US/10123262
; Publication No.: US20030049816A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C38
; CURRENT APPLICATION NUMBER: US/10/123,262
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-262-173

Alignment Scores:
Pred. No.: 4,2e-102 Length: 1210
Score: 848.00 Matches: 162
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-20 (1-162) x US-10-123-262-173 (1-1210)
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20
```



```

Db 603 CGGAGGACCCAGGATCTCCACTGTGGACCATCGAGGCTCTGGTGATGAACTAGAAATGG 662
QY 21 GLuIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 663 GAAATGGCCAGGTGGACCCCAAGAGACATTCAGATGGGATCTTCCGGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 723 GATGGCAGCCAGTCAGTGTGGAGGTGGCTTATGCCCTCAGAGGCCACCTCACAGAG 782
QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 783 CTGCTGGAGGAGATATGTGACCGGATGAAGGAGTGGGGAACAGATTGATCTTCCACC 842
QY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGlySerSerGluLeuAspLeu 100
Db 843 CATCGCAAGAACTACGTACGTGTACTGGCCGGAATGGAGAAATCCAGTGAATCGGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 903 CAAGGCATCCGAATCGACTCAGATATATAGCGCACCTCAAGTTTGGGTGAGAGCATTT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValys 140
Db 963 GTGGAGGAATACGAGGATGAACCTCAATTGAATCTTTCCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
Db 1023 GACAACCTTTCAGTAAAGCAAGACAGATCTTGTGTGACATGCCCTGCACATATCGCATGAT 1082
QY 161 GluLeu 162
Db 1083 GAGCTA 1088

```

Search completed: March 31, 2003, 11:59:11  
 Job time : 51.9767 secs

10

11

12

13

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 35.3198 Seconds  
(without alignments)  
1406.624 Million cell updates/sec

Title: US-10-082-502-20  
Sequence: 1 RRSQDLHCAGCALVDELEW.....LCSKRITDCHALHSHDEL 162

Scoring table: BLOSUM62  
Xgapop 10.0, Ygapext 0.5  
Fgapop 6.0, Ygapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=rlp  
-O=cgnt2.1/USPTO.spool/US10082502/rnuna.24032003.135100.6616/app\_query.fasta.1.1308  
-DB=Issued\_Patents\_NA\_OPM-fastap -SUFFIX=rl -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US10082502.ecgn.1.131.0runat.24032003.135100.6616 -NCPU=6 -ICPU=3  
-MRN.XLPHY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgnt2.6/ptodata/1/lna/5A.COMB.seq:\*  
2: /cgnt2.6/ptodata/1/lna/5B.COMB.seq:\*  
3: /cgnt2.6/ptodata/1/lna/6A.COMB.seq:\*  
4: /cgnt2.6/ptodata/1/lna/6B.COMB.seq:\*  
5: /cgnt2.6/ptodata/1/lna/PCITUS.COMB.seq:\*  
6: /cgnt2.6/ptodata/1/lna/Backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	15.2	744	4	US-09-163-285-3
2	129	15.2	1512	4	US-09-163-285-1
3	91.5	10.8	915	4	US-09-724-864-14
4	85	10.0	5181	1	US-08-257-073-10
5	81.5	9.6	1924	3	US-08-961-083-159
6	80.5	9.6	4394	2	US-08-750-152A-1
7	80.5	9.5	1216	4	US-09-071-035-3
8	80.5	9.5	1347	4	US-09-071-035-1
9	80.5	9.5	3752	4	US-08-961-527-208
10	77.5	9.1	2508	2	US-08-850-993-1
11	76.5	9.0	4865	3	US-08-894-017-24
12	75	8.8	5558	4	US-08-961-527-103

13	74.5	8.8	833	2	US-08-837-029-1	Sequence 1, Appl1
14	74.5	8.8	3372	1	US-07-906-349A-1	Sequence 1, Appl1
15	74.5	8.8	3372	1	US-08-167-035-1	Sequence 1, Appl1
16	74.5	8.8	3372	1	US-08-167-035-1	Sequence 48, Appl1
17	74.5	8.8	3372	1	US-08-208-887A-1	Sequence 1, Appl1
18	74.5	8.8	3372	2	US-08-539-005-1	Sequence 1, Appl1
19	74.5	8.8	3372	2	US-08-539-005-1	Sequence 48, Appl1
20	74.5	8.8	3372	3	US-09-344-521-1	Sequence 1, Appl1
21	74.5	8.8	3372	4	US-09-280-598-1	Sequence 1, Appl1
22	74.5	8.8	5177	6	5352450-1	Patent No. 5352450
23	74	8.7	5661	4	US-08-938-105-2	Sequence 2, Appl1
24	73.5	8.7	1776	1	US-08-522-229B-1	Sequence 1, Appl1
25	73.5	8.7	1776	2	US-09-102-433-1	Sequence 1, Appl1
26	73	8.6	2012	2	US-08-484-200-3	Sequence 3, Appl1
27	73	8.6	4131	1	US-07-872-644-38	Sequence 38, Appl1
28	73	8.6	4131	1	US-08-297-494-38	Sequence 38, Appl1
29	73	8.6	4131	1	US-08-297-510-38	Sequence 38, Appl1
30	73	8.6	4131	1	US-08-479-532-38	Sequence 38, Appl1
31	73	8.6	4131	1	US-08-455-526-38	Sequence 38, Appl1
32	73	8.6	4131	1	US-08-455-526-38	Sequence 38, Appl1
33	73	8.6	4131	1	US-09-139-491-38	Sequence 38, Appl1
34	73	8.6	4131	5	PCT-US92-03222-38	Sequence 38, Appl1
35	72.5	8.5	4371	1	US-08-803-973-1	Sequence 1, Appl1
36	72.5	8.5	4371	1	US-08-803-973-1	Sequence 1, Appl1
37	72	8.5	1640	4	US-08-676-444-41	Sequence 1130, Appl1
38	71.5	8.4	676	4	US-08-998-416-1130	Sequence 928, Appl1
39	71.5	8.4	724	4	US-08-998-416-928	Sequence 3, Appl1
40	71.5	8.4	41708	4	US-09-470-512A-3	Sequence 2101, Appl1
41	71	8.4	2073	4	US-09-134-001C-2101	Sequence 2, Appl1
42	71	8.4	2160	2	US-08-840-236-2	Sequence 2, Appl1
43	71	8.4	2160	2	US-08-840-236-5	Sequence 2, Appl1
44	71	8.4	2160	2	US-08-505-448A-2	Sequence 2, Appl1
45	71	8.4	2160	2	US-08-505-448A-5	Sequence 5, Appl1

## ALIGNMENTS

RESULT 1  
US-09-163-285-3  
Sequence 3, Application US/09163285  
Patent No. 6204013  
GENERAL INFORMATION:  
APPLICANT: Rhodadoust, Mehran  
TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,285  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/090,398  
FILING DATE: June 24, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandagouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MNI-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:



Alignment Scores:	
Pred. No.:	0.0112
Score:	91.50
Percent Similarity:	42.07%
Best Local Similarity:	24.39%
Query Match:	10.79%
DB:	4
Length:	915
Matches:	40
Conservative:	29
Mismatches:	66
Indels:	29
Gaps:	8

OY 6 LeuHISGSLVGLIACYSARGVALALEUVALAAspLLEuGLuTPrGLuLLEaLGLuVal 25  
 Db 158 CTGGCGTGCAGTGGCTGCCGGCGTGTGGCCCTTCACAGTGGGGCAAGCTGTGGCCAAAAGCA 217  
 OY 26 AspPrOLysLysThrIleGLuMetGLySerPheArgIleAsnProAspGLySerGlnSer 45  
 Db 218 GAGGCTAAATCTCAC-----ACTCCAGCGCCAGT---GGA 250  
 OY 46 ValValGLuValProTyrAlaArgSerGLuAlaHisLeuThrGluLeuLeuGLuGluIle 65  
 Db 251 TTGGAGAGAGTG-----AGGAATCCACGTCACAGATGTCTGGACCCAAAGC 298  
 OY 66 CysAspArg---MetLysGLuTyrGLyGLuGluIleAspProSerThrHisArgLysAsn 84  
 Db 299 TGGTCTCAGAACTGGCGAGTCTATGGA-----GTTATGAAGTGAAC 340  
 OY 85 TyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeuGlnGlyIleArg 104  
 Db 341 CAGATAGAGGTGTACAGGGCCACGAGCTTAGCAAGGGCCCAAG-----CCAAAG 391  
 OY 105 IleAspSerAspIleSerGly-----ThrLeuAspPheAlaCysGluSer 119  
 Db 392 ATCAGCGTGATGATTTCTGGGGGTCCCTGGCCCAATAGGCTCTCCAAAGAGCTTTTCCAC 451  
 OY 120 IleValGLuGLuTyr---GluAspGLuLeuIleGLuPhePheSerArgGLuAlaAspAsn 138  
 Db 452 TACCTGGGTGAGTTGTGGAGAGCAACGATCTATGAAGCCTACCGCCAAAGCCAAAGCAAT 511  
 OY 139 ValLysAspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSer 158  
 Db 512 CTGGAGAGCCGTGCTGTGGGGCCACCAGTGGCCCTGCTCACAGAGAGTCTCTGGGCCCA 571

RESULT 4  
 US-08-257-073-10  
 Sequence 10, Application US/08257073  
 Patent No. 5765597  
 GENERAL INFORMATION:  
 APPLICANT: Poletti, Enzo  
 APPLICANT: de Taisne, Charles  
 APPLICANT: Fine, John A.  
 TITLE OF INVENTION: MALARIA RECOMBINANT POXYVIRUS VACCINE

Alignment Scores:	
Pred. No.:	0.956
Score:	85.00
Percent Similarity:	43.31%
Best local Similarity:	22.93%
Query Match:	10.02%
DBs:	1
	7
Length:	5181
Matches:	36
Conservative:	32
Mismatches:	59
Indels:	30
Gaps:	7

QY 16 AspluIdueIuItrpGluIlealagIvalAspProLyIyThrIleIdmetGlySer 35  
 Db 1192 GATCACCCTTGATTAGATATTTTTRACAGAAAAAATTAATGATATTAACACT 1251  
 QY 36 PheArgIleAsnProAspGlySerGlnSerValaIgluValProTyrlaArgserGlu 55  
 Db 1252 AATATCAACAAATGCTTCGAAATGTGTTCAAATCCAAAAAGTTCTTATCCAAATGGTAT 1311  
 QY 56 AlaHis-----LeuThrGluIleuLeuGluIleGly-----AspArgmetGly 70  
 Db 1312 GTATATCCTTTACACGACTGATATTCATTAATTCATTAGCTGCAGATATATATAAAT 1371  
 QY 71 GluTyrlGlyIuIuInIleAspProSerThrHisArgLyAsnTyValaIrgValaIglu 90  
 Db 1372 TCATATGCTATTTAATGAATCCTGATACTAAAGAAAAATTAATGCAAAAAATTTATTACA 1431  
 QY 91 ArgAsnGlyGlu-----SerSerGluIleuAspIleuInGly 102  
 Db 1432 GATATATAGGAAGAAAAATATTCATTATATTAACATTAAAAAACAATGATTTTGACAA 1491  
 QY 103 IleArgIleAspSerAspIleSerGlyThrIleuLyIyPheAlaCysGluSerTlleValaIglu 122

Db 1492 AAAAAATATATAC-----ACAAAAAGACAAATAAAAAATTAATTGAA 1536  
Qy 123 GATGTCGlu-----AspGluLeuIleGlu--PhePheSerArgGlu 135  
Db 1537 GATTATGAAAGCAAAAAAGCATTTATGAGATTAAGTAAATTTTATGAAATCAAA 1596  
Qy 136 AlasAspAn-----VallysAspLysLeuGlySerLysArg 147  
Db 1597 TTTAAATTAATTAATTTTGACAAAGATGTCGTAGATTAATATTCAGTCAAGA 1647  
RESULT 5  
US-08-961-083-159  
Sequence 159, Application US/08961083  
Patent No. 6159469  
GENERAL INFORMATION:  
APPLICANT: ChOI et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,083  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 159:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1924 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-083-159  
Alignment Scores:  
Pred. No.: 0.655 Length: 1924  
Score: 81.50 Matches: 32  
Percent Similarity: 39.738 Conservative: 26  
Best Local Similarity: 21.928 Mismatches: 65  
Query Match: 9.61% Indels: 23  
Gaps: 4  
US-10-082-502-20 (1-162) x US-08-961-083-159 (1-1924)  
Qy 15 ValAspGluLeuGluTyrGluIleAla-----GlnValAspProLysLysThr 30  
Db 575 ATTCTACTTGTGGAACAAGATTGCTACTGCTCAACATCAAGTATGTTGAAAAAA 634  
Qy 31 IleGlnMeGlySerPheArgIleAsnProAspGlySerGlnSerValValGluValPro 50  
Db 635 CTCTCTTGCTGGTGG-----GATCCTGATGATGACACAGAGTTATAGAACTAAA 685  
Qy 51 TTTAAATGSerGluAlaHisLeuThrGluLeuLeuGluIleGlyAspArgMetLys 70

Db 686 TTTAAAAAGAGAGAGCTGAGTAAACGCTTAACACAGCTGATGACAAAAAACAACA 745  
Qy 71 GluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgValIleGly 90  
Db 746 GAACCTTAAAAACTCTGTAGACAGCTGATCCGAGTAAGCTCAGATGATTTGAT 805  
Qy 91 ArgAsnGlyLysSerSerGluLeuAspLeuGlnClyIleArgIleAspSerAspIleSer 110  
Db 806 AAGAGACAGACAGAGAGCTGATGATTAAGCTGATCAACTCAAAATAAGTTGCT 865  
Qy 111 GylThrLeuLysPheAlaCysGluSerIleValGluGluTyrGluAspGluLeuIleGlu 130  
Db 866 GATTTA-----VallysAspLysLeuGly 895  
Qy 131 PhePheSerArgGluAlaAspAn-----VallysAspLysLeuGly 144  
Db 896 ATATTACTTGAGGCGGTGATNCTGAAGATGATAGTCTGCTCTCTCAATAATTAATGCT 955  
Qy 145 SerLysArgThrAspLeu 150  
Db 956 ACTAAAAAGCTGAATTG 973  
RESULT 6  
US-08-750-152A-1  
Sequence 1, Application US/08750152A  
Patent No. 5977331  
GENERAL INFORMATION:  
APPLICANT: ASAKURA, YOKO  
APPLICANT: KIMURA, EICHIRO  
APPLICANT: ABE, CHIZU  
APPLICANT: KAWAHARA, YOSHIO  
APPLICANT: KANAMATSU, TSUYOSHI  
TITLE OF INVENTION: ALPHA-KETOGLUTARATE DEHYDROGENASE GENE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22152  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,152A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: NORMAN F. OBLON  
REGISTRATION NUMBER: 24,618  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4394 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: genomic DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Brevibacterium lactofermentum  
STRAIN: ATCC13869  
FEATURE:  
NAME/KEY: CDS







```

Oy      134  ArgGluLhlaSpAspEnValYs-----AspLYsLeuCYsSer 145
          |||||  ::::  |||||  |||||  |||
Db      1579  CGTAAAGCGCATGCGAAAGAAATATCAAAAGGATATGCGATATATGATATGATTAAGTTGAAGTCT 1638

Oy      146  LysArGThrasLeuCYsAspRh1sAlaLeuH1s1LeSerH1sAsp 160
          ::::  ::::  ::::  ::::  ::::  |||
Db      1639  CGATCGATGCAATATATTGACAGTACAGAAAGAAATGTGGAAAGAC 1683

RESULT 11
US-08-894-017-24
? Sequence 24, Application US/08894017
? Patent No. 6024958
? GENERAL INFORMATION:
? APPLICANT: Lehner, Thomas
? APPLICANT: Kelly, Charles
? TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE
? TITLE OF INVENTION: OF COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/III
? NUMBER OF SEQUENCES: 27
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: MORRISON & FOERSTER
? STREET: 2000 Pennsylvania Avenue, NW, suite 5500
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20006-1888
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows
? SOFTWARE: FASTSEQ for Windows Version 2.0b
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/894,017
? FILING DATE: 31-JUL-1997
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/GB96/00207
? FILING DATE: 31-JAN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 25150-20067.00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-887-1500
? TELEFAX: 202-822-0168
? TELEX:
? INFORMATION FOR SEQ ID NO: 24:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 4865 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: Genomic DNA
? US-08-894-017-24

Alignment Scores:
? Pred. No.: 11.2 Length: 4865
? Score: 76.50 Matches: 28
? Percent Similarity: 51.75% Conservative: 31
? Best Local Similarity: 24.56% Mismatches: 40
? Query Match: 9.02% Indels: 16
? DB: 3 Gaps: 6

US-10-082-502-20 (1-162) x US-08-894-017-24 (1-4865)
Oy      57  HisLeuThrgluLeuLeuGluGluLhlaCYsAspArgmetLysGluTYrgY----- 73
          |||  ::::  |||||  |||||  |||
Db      260  CACAACACGGAATTCGACCGCACCAATTGCGCAGAGGCTCAAGAAATGCGAGATGAACCAAG 319

Oy      74  ---GluGlnIleAspProSerThrHisArgLYs-----AsnTYValArg 87
          ::::  |||||  |||  |||  ::::  ::::
Db      320  CTGAACAACAGTCAACCAACAGCTGAGAGACAAATAGTTCATTACCATTGAAGTACCTTAAA 379

```

QY 88 ValValGlyArgAnGlyGlySerSerGluLeuAspLeuGlnGlyIleArgIle----- 105  
Db 380 CTGATTC---TTGATCAGACAGCAAAAGCATGCTAGTCTGCT-GGTGTCATGTTGTCCAA 435  
QY 106 AspSerAspIleSer---GlyThrLeuLysPheAlaCysGlySerIleValGluGlyTyr 124  
Db 436 GATGCCATCTTATTAAGAACTGTTAAACAGCTCAAGAACAGCAGC-----CAAAA 489  
QY 125 GluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLysAspLysLeuCys 144  
Db 490 GAAACGTAAATTAAGAAAGATTAACACAAACAGCTGAGATATTAAAGAGAACAGAT 549  
QY 145 SerLysArgThrAspLeuCysAspHisAlaLeuHisIleSer 158  
Db 550 CAATATTAATCGATGTAGCTGCTCATGAGCAGAAAGTTGCT 591  
RESULT 12  
US-08-961-527-103  
Sequence 103, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5558 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-103  
Alignment Scores:  
Pred. No.: 21.2 Length: 5558  
Score: 75.00 Matches: 36  
Percent Similarity: 37.85% Conservative: 31  
Best Local Similarity: 20.348 Mismatches: 62  
Query Match: 8.84% Indels: 48  
Gaps: 8  
US-10-082-502-20 (1-162) x US-08-961-527-103 (1-5558)  
QY 14 LeuValAspGluLeu-----GluTrpGluIleAlaGlnValAspProLys 28  
Db 2094 ATGATAGACAAATATTAATACGTCATTTGAGAGATGGAATTCCTCGGAATGAGAACAAA 2153  
QY 29 LysThrIleGlnMetGlySerPheArgIleAsnProAspLysSerGlnSerValValGlu 48

Db 2154 GGCATTCTCATAGTTTGGAGATAGCCCTTGACCAATCATATGCTACACATCCAAATA 2213  
QY 49 ValProTyraIlaArgSerGluAlaHisLeuThrGluLeuLeuGlnGluIleCysAspArg 68  
Db 2214 GTTGGTGGCTGCGCCAAACAGCAAGATATACAGCTTTTCAAGAGAAAGTTTGC----- 2267  
QY 69 MetLysGluTyrGlyGlnGlnIleAspProSerThrIlaArgLysAsnTyrValArgVal 88  
Db 2268 CAAGAAGACCTTGCTGTTCTTGG-----CCACCTGACAAAGAAAGTTATCATCCGATAC 2318  
QY 89 ValGlyArg-----AsnGlyGlySerSer-----GluLeuAspLeuGlnGly 102  
Db 2319 CAGGCCAAAGTCCATTAACACCAAGACGCGCTACTTCTGTCATATCTTACATCCAAAGT 2378  
QY 103 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlySerIleValGlu 122  
Db 2379 ATGAAATC---ACGACTCTCCAGACGGCTTGAAGTTCTCC----- 2417  
QY 123 GluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys----- 140  
Db 2418 -----TACTTCTTCATGAGACATCAACATCCGGCGGCTC 2453  
QY 141 -----AspLysLeuCysSer 145  
Db 2454 TTCAGCCATTTTCATATAGAGTCATTGATACAGCTTCAGCTTTGAAAGTCATCAAAA 2513  
QY 146 LysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAspGluLeu 162  
Db 2514 AGCCGTAGGAGAAACATCAGCACCGACGCTCTTTCAGCAAGACAGAGTG 2564  
RESULT 13  
US-08-837-029-1  
Sequence 1, Application US/08837029  
Patent No. 5945303  
GENERAL INFORMATION:  
APPLICANT: Wei et al.  
TITLE OF INVENTION: Human Hematopoietic - Specific Protein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/837,029  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US96/04930  
FILING DATE: 11-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF268  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 833 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:

NAME/KEY: CDS  
LOCATION: 42..608  
FEATURE: sig\_peptide  
LOCATION: 42..107  
FEATURE: mat\_peptide  
LOCATION: 108..608  
US-08-837-029-1

Alignment Scores:  
Pred. No.: 1.6  
Score: 74.50  
Percent Similarity: 39.13%  
Best Local Similarity: 21.12%  
Query Match: 8.79%  
Length: 833  
Matches: 34  
Conservative: 29  
Mismatch: 75  
Indels: 23  
Gaps: 7

US-10-082-502-20 (1-162) x US-08-837-029-1 (1-833)

QY 6 LeuHlaCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrpGluIleAlaGlnVal 25  
DB 183 CTGGCGCTGATGCTGACAGAGCTGCTTACCAGATGGCAAAATCTGCCAAAGGCA 242  
QY 26 AspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer 45  
DB 243 GAGACCAAACTGATACCTCA-----AACTCTGGGGGGCGCGGCA 284  
QY 46 ValValGluValProGlyAlaArgSerGluAlaHisLeuThrGluLeuGluGlnIle 65  
DB 285 CTAGCGAGTGGTGGTAC-----ACGGATGCTCTGGACCGGCAAC 323  
QY 66 CysAspArg---MetLysGluTrpGly---GluGlnIleAsnProSerThrHisArgLys 83  
DB 324 TGCTCCCGGAACTGGAGAGGACTACGAGATTCGAAAGTGAC----- 365  
QY 84 AsnTyValArgValIleGlyArgAsn---GlyLysSerSerGluLeuAspLeuGlnGly 102  
DB 366 CAAGTAAACGCTCTCAGAGGCCGAGACTAGCAGGGGCCAGAGCCAGCATCAGCGTC 425  
QY 103 IleArgIleAsnSerAspIleSerGlyThrLysLysPheAlaCysGluSerIleValGlu 122  
DB 426 ATGGTACAGGGGGCCCTGGCTCCAGGCTCTCCAGGACATGTTGGACTTGGGG 485  
QY 123 GluTrp---GluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAsp 141  
DB 486 GAGTTGGAGAGACCAAGATGTATGAAAGCCCAAGAGCCGAGGGCTTGGAGCA 545  
QY 142 LysLeuCysSerLysArgThrAspLysCysAsnHisAlaLeuHisIleSerHisArgLys 161  
DB 546 TTGCTATGTGGGGGAGCCGAGGGGGCTCTCAGAGAGGTGTCAAGCCCAAGAGAGAG 605  
QY 162 Leu 162  
DB 606 CTC 608

## RESULT 14

US-07-906-349A-1  
Sequence 1, Application US/07906349A  
Patent No. 5434064  
GENERAL INFORMATION:  
APPLICANT: Schlessinger, Joseph  
APPLICANT: Skolnik, Edward Y.  
APPLICANT: Margolis, Benjamin L.  
TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR  
IDENTIFYING TARGET PROTEINS FOR ENDOCYTIC TYROSINE KINASES AND  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Broadway and Nelmark  
STREET: 419 Seventh Street, N.W.  
CITY: Washington  
STATE: D.C.

COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/906,349A  
FILING DATE: 30-JUN-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/643,237  
FILING DATE: 18-JAN-1991  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3372 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-906-349A-1

Alignment Scores:  
Pred. No.: 12  
Score: 74.50  
Percent Similarity: 40.65%  
Best Local Similarity: 20.65%  
Query Match: 8.79%  
Length: 3372  
Matches: 32  
Conservative: 31  
Mismatch: 57  
Indels: 35  
Gaps: 8

US-10-082-502-20 (1-162) x US-07-906-349A-1 (1-3372)

QY 22 IleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnProAsp 41  
DB 1279 CTAGCTGATGATATGAAATCCCAATGGATGTGAATTA-----CTTATTCAGTA 1326  
QY 42 GlySerGlnSerValGluValProGlyAlaArgSerGluAlaHisLeuThrGluLeu 61  
DB 1327 TCCAAA-----TACCAAGAGATCAAGTGTCCAAAGAGATAT 1365  
QY 62 LeuGluGlnIleCysAspArgMetLysGluTrpGlyGluGlnIleAsnProSerThrHis 81  
DB 1366 ATTGAAGCTGTAGGAAATAATACATGATTAACACTGTTCAAGAAAAAGTCA 1425  
QY 82 Arg-----LysAsnTyValArgValIleGlyArgAsnGlyLysSer 96  
DB 1426 GAATATGATGATATATGAAATATATACCCGAC-----TCCAG 1467  
QY 97 GluLeuAspLeuGlnGlyIleArgIleAsnSerAspIleSerGlyThrLysPhe 115  
DB 1468 GAATCCAAATGAAAGAGACGATTTGAAGCA---TTTATGAAACCAATAAAATATTT 1524  
QY 116 ---AlaCysGluSerIleValGluGluTrpGluAspLeuIleGluPheSer 133  
DB 1525 GAGACAGCTGCCGAGCC---CAAGAGCGGTACAGAAAGATCATAGAAAGTTAA 1581  
QY 134 ArgGluAlaAspAsnValLys-----AspLysLeuCysSer 145  
DB 1582 CGTAAAGCAATGAAAGAAATTCAAAGATTTATGATATGATATGATGATGATGATGAT 1641  
QY 146 LysArgThrAspLysCysAsnHisAlaLeuHisIleSerHisArgLys 160  
DB 1642 CGAATCAGTGAATTTATTTGACAGTAGAGAAAGATTTGGAAGAAC 1686

## RESULT 15

US-08-167-035-1  
Sequence 1, Application US/08167035  
Patent No. 5618691  
GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph  
 APPLICANT: Skolnick, Edward Y.  
 APPLICANT: Margolis, Benjamin L.  
 TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR  
 TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR ENKAROTIC TYROSINE  
 TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PENNIE & EDMONDS  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: 10036-2711  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/167,035  
 FILING DATE: 16-DEC-1993  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Coruzzi, Laura A.  
 REGISTRATION NUMBER: 30,742  
 REFERENCE/DOCKET NUMBER: 7683-062  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 790-9090  
 TELEFAX: (212) 869-9741/8864  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3372 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: unknown  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 43..2214  
 US-08-167-035-1  
 Alignment Scores:  
 Pred. No.: 12 Length: 3372  
 Score: 74.50 Matches: 32  
 Percent Similarity: 40.65% Conservative: 31  
 Best Local Similarity: 20.65% Mismatches: 57  
 Query Match: 8.79% Indels: 35  
 DB: 1 Gaps: 8  
 US-10-082-502-20 (1-162) x US-08-167-035-1 (1-3372)  
 QY 22 ILAALAGLVALASPPOLYSLVSTHLLGLMELCLYSERPHARGILEASPROASP 41  
 DB 1279 CTAGCTAGTATATATCCCAATTGATGTAATA-----CTTATCCAGTA 1326  
 QY 42 GlycerInserValValGluValProTyrAlaArgSerGluAlaHisLeuThrGluLeu 61  
 DB 1327 TCCAAA-----TACCACACGATCAAGTGTCCAAAGAAAGATAT 1365  
 QY 62 LeuGluGluIleCysAspArgMetLysGluTyrGluGluGluIleAspProSerThrHis 81  
 DB 1366 ATTGAAGCTGTAGCAAAATAATACATGATATACACTGCTTCAAGAAAAAGTCGA 1425  
 QY 82 Arg-----LysAsnTyrValArgValValGluArgAsnGlyGluSerSer 96  
 DB 1426 GAATATGATATGATATATGTAAGATATATACCGACACA-----TCCAG 1467  
 QY 97 GluLeuAspLeuGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPhe--- 115  
 DB 1468 GAATATCAATGAAAGACAGTATGAGCA---TTTAATGAACCATATAATATT 1524

QY 116 -----AlaCysGluSerIleValGluGluTyrGluAspGluLeuIleGluPhePheSer 133  
 DB 1525 GAGAACAGTGCACAGCC---CAAGACGGTACAGCAAAAGATACATAGAAAGTTTAAA 1581  
 QY 134 ArgGluAlaAspAsnValLys-----AspLysLeuCysSer 145  
 DB 1582 CGTGAAGCAATGAGAAAGAAATACAAAGATTATGCAATATATGTAAGTGAAGCT 1641  
 QY 146 LysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 1642 GAATATCAATGAAATATATGACAGTGAAGAAAGANTTGAAGACAGC 1686

Search completed: March 31, 2003, 04:26:11  
 Job time : 42.3198 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:08:03 ; Search time 127.151 Seconds

(without alignments)  
2869, 213 Million cell updates/sec

Title: US-10-082-502-20

Perfect score: 848  
Sequence: 1 RRSODLHCACRALVDELEW.....LCSKRTDLCDAHLSHDEL 162

Scoring table:

BLOSUM62  
Xgapexpt 0.5  
Ygapop 10.0, Ygapext 0.5  
Ygapop 6.0, Ygapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ p2n.model -DEV-xml  
-O/cgn2.1/USPTO.spool/US10082502/runat\_24032003.135058.6538/app.query.fasta\_1.1308  
-DB-n\_Geneseq\_101002 -QFWT-fastexp -SUFFIX-rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR-SCORE-pct -THR-MAX=100 -THR-MIN=0 -ALIGN=15  
-MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10082502.CCGN.1.1.741 &runat\_24032003.135058.6538 -NCPU=6 -ICPU=3  
-NO\_XLPY -NO\_MMAP -LARGEQUERY -NEG-SCORE=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-MARR-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	848	100.0	546	AAZ38326	Human transmembran
2	848	100.0	806	AAZ06969	Human secretory pe
3	848	100.0	806	AAZ08293	Human Zs19g gene e
4	848	100.0	814	AAZ38327	Human transmembran
5	848	100.0	814	ABK52765	cDNA encoding tran
6	848	100.0	814	ABL41995	Nucleotide sequenc
7	848	100.0	814	ABK09772	Human ovarian tumo
8	848	100.0	1180	AAZ22112	Human secreted pro
9	848	100.0	1210	AAZ21330	Human cDNA sequenc
10	841	99.2	714	20 AAX7884	Human secreted pro
11	830	97.9	1069	20 AAX06970	Human secretory pe
12	830	97.9	1069	21 AAZ08294	Mouse secretory pe
13	829	97.8	657	24 ABK36007	Mouse ortholog gen
14	799	94.2	832	20 AAX97837	Human secreted pro
15	746.5	88.0	592	21 AAA43525	Human secreted exp
16	393	46.3	484	22 AAL25628	Human breast cance
17	393	46.3	558	22 AAL18023	Human breast cance
18	393	46.3	595	22 AAL07882	Human breast cance
19	393	46.3	1160	22 AAL26689	Human breast cance
20	257	30.3	649	20 AAX97883	Human secreted pro
21	257	30.3	649	20 AAX06968	Human secretory pe
22	257	30.3	649	21 AAZ08284	Human Zs19g gene.
23	257	30.3	718	20 AAZ10655	cDNA encoding a hu
24	254	30.0	792	20 AAX97836	Human secreted pro
25	253	29.8	415	20 AAX06971	Zs19g expressed se
26	253	29.8	415	21 AAZ08285	Expressed Sequence
27	243	28.7	1085	20 AAX22130	WO9901020 Seg ID 3
28	205	24.2	2532	22 ABA07341	Human pancreatic c
29	205	24.2	2532	22 AAS33768	Human genomic DNA
30	205	24.2	2724	22 ABA07340	Human pancreatic c
31	205	24.2	2724	22 AAS33767	Human genomic DNA
32	152	18.2	816	23 ABL21471	Drosophila melanog
33	152	18.2	564	24 ABK09691	Human ovarian tumo
34	154	18.2	786	21 AAX37767	Human secreted pro
35	141.5	16.7	750	23 ABL16485	Arabidopsis thalia
36	141.5	16.7	2750	23 ABL16484	Drosophila melanog
37	141.5	16.7	3274	23 ABL16482	Drosophila melanog
38	141	16.6	3110	23 ABL21470	Drosophila melanog
39	129	15.2	744	21 AAZ49678	Human myocardium s
40	129	15.2	746	22 AAF94104	Primer specific fo
41	129	15.2	747	22 AAC90704	Human secretory pr
42	129	15.2	1419	22 AAF93914	Human cDNA encodin
43	129	15.2	1457	20 AAX78923	Human zcalc-1 DNA
44	129	15.2	1462	21 AAA16621	Human secreted pro
45	129	15.2	1470	21 AAA96349	cDNA encoding a no

## ALIGNMENTS

RESULT 1  
AAZ38326  
AAZ38326 standard; cDNA: 546 BP.

AAZ38326;

09-FEB-2000 (first entry)

Human transmembrane protein cDNA clone HP10390 coding sequence.

HP10390; transmembrane domain; stomach cancer cell; antibody;

assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;

agonist; antagonist; ligand; therapeutic; ds.

OS Homo sapiens.  
XX key  
XX location/Qualifiers  
XX 1.540  
FT CDS

```

FT      /*tag- a      /product- "Human transmembrane protein HP10390"
FT      /note- "No stop codon given in the specification"
XX      MO9955862-A2.
XX      04-NOV-1999.
XX      27-APR-1999; 99WO-JP02226.
XX      28-APR-1998; 98JP-0119395.
XX      (SAGA ) SAGAMI CHEM RES CENT.
XX      (PROT-) PROTEGENE INC.
XX      Kato S, Kimura T;
XX      WPI: 2000-023358/02.
XX      P-PSDB; AAY52391.
XX      Human proteins with transmembrane domains, involved in control of cell
XX      proliferation and differentiation, useful for treating e.g. cancer or
XX      inflammation
XX      Claim 3; Page 88; 114pp; English.
XX      This sequence represents the coding sequence of human CDNA clone
XX      HP10390 which encodes a 20 kD protein with one putative transmembrane
XX      domain in the N-terminus. The CDNA was isolated from a human stomach
XX      cancer cell line CDNA library. The protein has no homology with any
XX      known protein. The protein may be used to raise specific antibodies, as
XX      assay reagents, as diagnostic tissue markers, for the isolation of
XX      cognate receptors, ligands and binding proteins, and as biologically
XX      active agents. Nucleotides encoding the protein may be used as primers
XX      and probes or antisense molecules, and in gene therapy. Cells transformed
XX      with these nucleotides may be used to screen for agonists and antagonists
XX      which are potentially useful therapeutically.
XX      Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;
XX      SO
XX      Alignment Scores:
XX      Pred. No.: 2,2e-94 Length: 546
XX      Score: 848.00 Matches: 162
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 21 Gaps: 0
XX      US-10-082-502-20 (1-162) x AAZ38326 (1-546)
XX      OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
XX      |||||||
XX      DB 61 CGGAGGAGCGAGATCTCCACTGTGAGCATGCGAGGCTGTGGATGAGTAATGATGG 120
XX      |||||||
XX      OY 21 GluIleAlaGlnValAspProLysThrIleGlnMetLysSerPheAsnGluLeuAspPro 40
XX      |||||||
XX      DB 121 GAAATTCGCCAGGAGGAGCCCAAGAACACATTCAGATGGATCTTCCGATCAATCA 180
XX      |||||||
XX      OY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
XX      |||||||
XX      DB 181 GATGGCAGCCAGTCAAGTGGAGGAGTCCCTTATGCCCGCTCAAGGCCCACTCAGAG 240
XX      |||||||
XX      OY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
XX      |||||||
XX      DB 241 CTCCTGAGAGAGATATGTGACCGAGTGAAGAGATATGGGAAACAGATTATCTTCAC 300
XX      |||||||
XX      OY 81 HisArgLysAsnTyrValAlaGlyValArgAsnGlyLysSerGluLeuAspLeu 100
XX      |||||||
XX      DB 301 CATCGGAAGAACTACATCTGTAGTGGCCGGAATGGAAATCCACTGAACTGGACCTA 360
XX      |||||||
XX      OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
XX      |||||||
XX      DB 361 CAAGGATCGAATTCAGTCAATATTAGCGGACCCCTCAAGTTTGCTGTGAGAGCATTT 420

```

```

OY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
XX      |||||||
DB 421 GTGAGCAATATCAGAGATGAGTCAATTCATTTCCGAGAGGCTGACAAATGTTAAA 480
XX      |||||||
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisLysSerHisAsp 160
XX      |||||||
DB 481 GACAAACTTTCAGTGAAGCGAAGATCTTGTGACCATGCTGCACATATGCAATGAT 540
XX      |||||||
OY 161 GluLeu 162
XX      |||||||
DB 541 GAGCTA 546
XX      |||||||
XX      RESULT 2
XX      ID AAX06969 standard; cDNA; 806 BP.
XX      AC AAX06969;
XX      XX 10-MAY-1999 (first entry)
XX      DT 10-MAY-1999 (first entry)
XX      DE Human secretory peptide-9 (Zs1g9) variant cDNA.
XX      KW Secretory peptide-9; Zs1g9; human; tumour marker; cancer; therapy;
XX      KW diagnosis; growth enhancer; variant; ss.
XX      OS Homo sapiens.
XX      FH Key Location/Qualifiers
XX      FT CDS 104..652
XX      FT sig_peptide 104..163
XX      FT mat_peptide 164..649
XX      FT /*tag- b
XX      FT /*tag- c
XX      PN W09901554-A1.
XX      PD 14-JAN-1999.
XX      PF 02-JUL-1998; 98WO-US13859.
XX      PR 17-JUN-1998; 98US-0099005.
XX      PR 03-JUL-1997; 97US-0051704.
XX      PR 03-JUL-1997; 97US-0888088.
XX      PR 19-MAY-1998; 98US-0081338.
XX      PR 19-MAY-1998; 98US-0085983.
XX      PR 17-JUN-1998; 98US-0089899.
XX      PA (ZYMO ) ZYMOGENETICS INC.
XX      PI Jaspers SR, Jellinek LJ, Shepard PO, Whitmore TE;
XX      PT WPI: 1999-106055/09.
XX      PT P-PSDB; AAM88474.
XX      PS Claim 2; Page 73-74; 85pp; English.
XX      CC This CDNA clone encodes human secretory peptide-9, or Zs1g9,
XX      CC variant (see AAM88474). Zs1g9 (see also AAM88469) is overexpressed in
XX      CC human brain, liver, lung, oesophageal, stomach, colon, rectal,
XX      CC thyroid and lymphoma tumors. Thus, Zs1g can be used as an
XX      CC indicator for cancer. Zs1g9 cDNA was discovered in a placenta
XX      CC clone from a full-term pregnancy cDNA library which contained an
XX      CC expressed sequence tag (see AAX06971). The invention provides
XX      CC polynucleotides (see AAX06968-70) encoding Zs1g9 polypeptides (see
XX      CC AAM88469-77) including mature polypeptides, other processed forms,
XX      CC variants and mouse orthologues. The Zs1g9 gene, or probes derived
XX      CC from it, can be used to determine if Zs1g9 is present on chromosome

```





Oy	61	LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr	80
Db	344	CTGCTGGAGGAGATATGTGACCGGATGAAGAGATATGCGGAGAACATTTATCTTCCACC	403
Oy	81	HisArgLysAsnTyrValAlaGValValGlyArgAsnGlyGluSerSerGluLeuAspLeu	100
Db	404	CATGGCAAGAACTACGTACCTGTATGGTGGCGGAATGGAGAAATCCAGTGAACCTGGACCTA	463
Oy	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
Db	464	CAGGCAATCCGAAATCGACTGAGATATTAGGGGCACTCAAGTTGCGTGTAGAGACATT	523
Oy	121	ValGluGluTyrGluAspArgGluLeuIleGluPheSerArgGluValAspAsnValLys	140
Db	524	GTGGAGGAATACGAGGATGAACTCATGTAATCTTTTCCGAGAGAGCTGACAAATTMAA	583
Oy	141	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp	160
Db	584	GACAAACTTTGGCACTAAGCGAACAGATCTTTGTGACCAATGCCCTGCACATATCGCATGAT	643
Oy	161	GluLeu 162	
Db	644	GAGCTA 649	
RESULT 4			
ID	AAZ38327	standard; cDNA; 814 BP.	
XX	AAZ38327;		
AC			
XX	AAZ38327;		
DT	09-FEB-2000	(first entry)	
XX			
DE	Human transmembrane protein cDNA clone HP10390.		
XX			
KW	HP10390; transmembrane domain; stomach cancer cell; antibody;		
KW	assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;		
KW	agonist; antagonist; ligand; therapeutic; ds.		
XX			
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	145..693	
FT		/*tag= a	
FT		/product= "human transmembrane protein HP10390"	
XX			
PN	W09955862-A2.		
XX			
PD	04-NOV-1999.		
XX			
PE	27-APR-1999;	99MO-JP02226.	
XX			
PR	28-APR-1998;	98UP-0119395.	
XX			
PA	(SAGA ) SAGAMI CHEM RES CENT.		
XX	(PROT-) PROTEGENE INC.		
XX			
PI	Kato S, Kimura T;		
XX			
DR	WPI: 2000-023358/02.		
DR	P-PSDB; AAY52391.		
XX			
PT	Human proteins with transmembrane domains, involved in control of cell		
PT	proliferation and differentiation, useful for treating e.g. cancer or		
PT	inflammation		
PS	Claim 4; Page 106-107; 114pp; English.		
XX			
XX			
CC	This sequence represents the human cDNA clone HP10390		
CC	which encodes a 20 kD protein with one putative transmembrane		
CC	domain in the N-terminus. The cDNA was isolated from a human stomach		
CC	cancer cell line cDNA library. The protein has no homology with any		
CC	known proteins. The protein may be used to raise specific antibodies, as		
CC	assay reagents, as diagnostic tissue markers, for the isolation of		

CC	cognitive receptors, ligands and binding proteins, and as biologically
CC	active agents. Nucleotides encoding the protein may be used as primers
CC	and probes or antisense molecules, and in gene therapy. Cells transformed
CC	with these nucleotides may be used to screen for agonists and antagonists
CC	which are potentially useful therapeutically.
XX	
SQ	Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;
Alignment Scores:	
Pred. NO.:	3,74e-94 Length: 814
Score:	848.00 Matches: 162
Percent Similarity:	100.00% Conservatave: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	21 Gaps: 0
US-10-082-502-20 (1-162) x AAZ3827 (1-814)	
OY	1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuThrGlu 20
Db	205 CGAGGAGGCCAGGATCTCCACTGTGGAGCATGACGGGCTGTGTGGAGAAGTAAGG 264
OY	21 GluIleAclnValAspProLysIleThrIleGlnMetGlySerPheargIleasPro 40
Db	265 GAATTCTCCAGTGGCACCCCAAGAACCAATTCAGATGGATCTTTCCGATCAATCCA 324
OY	41 AspArgSerGlnSerValValGluValProTyrrAlaArgSerGluValHisLeuThrGlu 60
Db	325 GATGGACGCCAGTCAGTGCTGGAGGCGCTTATGCCGCTCAGAGGCCCACTCACAAG 384
OY	61 LeuLeuGluGluIleCysAsparGmetLysGluTrgLyGluGlnIleaspProSerThr 80
Db	385 CTCTGGAGAGATATGTGACCAGGAAGAGGATAGGGGAACAATGATTCCTCCACC 444
OY	81 HisArgLysAsnTrpValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu 100
Db	445 CATGCCAAGAACTACGTACTGTAGTGGCCGGATGGAAATCCAGTAGAACCTGACCTA 504
OY	101 GlnGlyIleargIleasPserAspIleSerGlyThrIleuLysPheAlaCysgluSerIle 120
Db	505 CAAGGCATCCGAATCGACTCAGATATTACGGCACCCCTCAATTTGGCGTGAAGCATT 564
OY	121 ValGlnGluTrgIleuAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db	565 GTGGAGGAATACAGAGGTGAATCATTTGAATCTTTTCCCGAGAGGCTGACAAATGTTAA 624
OY	141 AsprLysLeucySserLysArgThrAspLeuCyAspRhisAlaLeuHisIleSerHisAsp 160
Db	625 GACAAACTTTGCAAGTAAGCAACAGATCTTTTGACACCAACCCCTGCACATATCCGATAT 684
OY	161 GluLeu 162
Db	685 GAGCTA 690
RESULT 5	
ID	ABK52765 standard; cDNA; 814 BP.
XX	ABK52765;
AC	
XX	
DT	15-AUG-2002 (first entry)
XX	
DE	cdNA encoding transmembrane protein 4, a cancer-linked protein.
KX	Expressed sequence tag; EST; human; cancer; anti-neoplastic;
KW	Cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;
KM	transmembrane protein 4; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 145..693
TT	/tag - a



CC purifying the antibody from the animal. The antibodies can be used  
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The  
CC present sequence encodes a polypeptide, designated HPI0390, which was  
CC used in the course of the invention.

XX  
SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other:

## Alignment Scores:

Pred. No.:	3,74e-94	Length:	814
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-082-502-20 (1-162) x ABL41995 (1-814)

```
OY 1 ArgArgSerGlnAspLeuHIScyGlyAlaCysArgAlaLeuValaAspGluLeuGluTrp 20
    |||||||
DB 205 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGGCTCTGTGGATGAACCTGAAATGG 264

OY 21 GluIleAlaGlnValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||||
DB 265 GAATTTGCCAGGTGGACCCCAAGAAACCATTCAGATGGATCTTCCGGATCAATCCA 324

OY 41 AspGlySerGlnSerValaIleGluValaProTyraIlaArgSerGluAlaHisLeuThrGlu 60
    |||||||
DB 325 GATGGCAGCCAGTCAGTGGTGAGAGTCCCTTATGCCGCTCAGAGGCCACCTCAGAGAG 384

OY 61 LeuLeuGluIleGlyAspArgMetLysGluTrpGlyGlnIleAspProSerThr 80
    |||||||
DB 385 CTGCTGAGGAGATATGTGACCGGATGAAGAGATAGGGGAACAGATTGATCTTCCACC 444

OY 81 HisArgLysAsnTyraIaArgValaIleGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100
    |||||||
DB 445 CATGGCAAGAACTGACGTACGTGTAGTGGCGGAGATGGAGATCCAGTAATGGACCTA 504

OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||||
DB 505 CAAAGCATCCGAATCGACATCGATATTACGCGCACCTCAAGTTGGGTGGAGAGCATTT 564

OY 121 ValGluGluTyraGlnAspGluLeuIleGluPheSerArgIleValaAspAsnValLys 140
    |||||||
DB 565 GTGAGAGAAATACGAGGATGAACCTATTGATTTCCCGAGAGGCTGACAAATGTTAAA 624

OY 141 AspLysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisIleSerHisAsp 160
    |||||||
DB 625 GACAACTTGCACTAGTACGAAACAGATCTTGTGACCATGCGCTGACATATCCATGAT 684

OY 161 GluLeu 162
    |||||||
DB 685 GAGCTA 690
```

## RESULT 7

ABK09772  
ID ABK09772 standard; cDNA: 814 BP.

XX AC ABK09772;

XX DT 14-MAR-2002 (first entry)

XX DE Human ovarian tumour protein encoding cDNA #305.

XX KW Human; ovarian tumour protein; cancer; cytosolic; immunostimulant; ss;

XX OS gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.

XX OS Homo sapiens.

XX PN WO200190154-A2.

XX PD 29-NOV-2001.

XX PF 23-MAY-2001; 2001WO-US16895.

XX 24-MAY-2000; 2000US-207107P.  
PR 13-JUN-2000; 2000US-211457P.  
PR 21-JUN-2000; 2000US-213673P.  
PR 03-AUG-2000; 2000US-223288P.  
PR 01-MAR-2001; 2001US-272790P.

XX (CORI-) CORIXA CORP.

XX Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MJ;

PI Algate PA, Fling SP, Mannion J, Benson DR, Carter D;

DR WPI; 2002-097641/13.

PT New isolated polynucleotide encoding polypeptide comprising portion of

PS human ovarian cancer

PS Claim 1; Page 253-254; 285pp; English.

CC The invention relates to an isolated polynucleotide encoding a  
CC polypeptide comprising a portion of an ovarian tumour protein. The  
CC sequences of the invention are useful for stimulating an immune response  
CC and for treating ovarian cancer in a patient. An antigen presenting cell  
CC that expresses the sequences is useful for treating ovarian cancer by  
CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells  
CC can then be proliferated and administered to the patient to inhibit the  
CC development of cancer. The DNA sequences are useful as probes or primers  
CC for nucleic acid hybridisation, to direct expression of a polypeptide in  
CC appropriate host cells, detecting the presence of a cancer in a patient  
CC involves obtaining a biological sample from the patient, contacting the  
CC biological sample with an agent that binds to the protein, detecting the  
CC amount of protein that binds to the agent, comparing the amount of  
CC protein to a predetermined cut-off value and determining the presence of  
CC cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA  
CC molecules encoding ovarian tumour proteins of the invention.

XX SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other:

## Alignment Scores:

Pred. No.:	3,74e-94	Length:	814
Score:	848.00	Matches:	162
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-082-502-20 (1-162) x ABK09772 (1-814)

```
OY 1 ArgArgSerGlnAspLeuHIScyGlyAlaCysArgAlaLeuValaAspGluLeuGluTrp 20
    |||||||
DB 205 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGGCTCTGTGGATGAACCTGAAATGG 264

OY 21 GluIleAlaGlnValaAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
    |||||||
DB 265 GAATTTGCCAGGTGGACCCCAAGAAACCATTCAGATGGATCTTCCGGATCAATCCA 324

OY 41 AspGlySerGlnSerValaIleGluValaProTyraIlaArgSerGluAlaHisLeuThrGlu 60
    |||||||
DB 325 GATGGCAGCCAGTCAGTGGTGAGAGTCCCTTATGCCGCTCAGAGGCCACCTCAGAGAG 384

OY 61 LeuLeuGluIleGlyAspArgMetLysGluTrpGlyGlnIleAspProSerThr 80
    |||||||
DB 385 CTGCTGAGGAGATATGTGACCGGATGAAGAGATAGGGGAACAGATTGATCTTCCACC 444

OY 81 HisArgLysAsnTyraIaArgValaIleGlyArgAsnGlyGlnSerSerGluLeuAspLeu 100
    |||||||
DB 445 CATGGCAAGAACTGACGTACGTGTAGTGGCGGAGATGGAGATCCAGTAATGGACCTA 504

OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||||
DB 505 CAAAGCATCCGAATCGACATCGATATTACGCGCACCTCAAGTTGGGTGGAGAGCATTT 564
```



PN HQ200140466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000MO-US32678.  
 XX  
 PR 01-DEC-1999; 99MO-US28301.  
 PR 01-DEC-1999; 99MO-US28634.  
 PR 02-DEC-1999; 99MO-US28551.  
 PR 02-DEC-1999; 99MO-US28551.  
 PR 02-DEC-1999; 99MO-US28565.  
 PR 09-DEC-1999; 99MO-US170262.  
 PR 16-DEC-1999; 99MO-US30095.  
 PR 20-DEC-1999; 99MO-US30911.  
 PR 20-DEC-1999; 99MO-US30999.  
 PR 30-DEC-1999; 99MO-US31243.  
 PR 06-JAN-2000; 2000MO-US00277.  
 PR 06-JAN-2000; 2000MO-US00376.  
 PR 11-FEB-2000; 2000MO-US03565.  
 PR 18-FEB-2000; 2000MO-US04341.  
 PR 18-FEB-2000; 2000MO-US04342.  
 PR 22-FEB-2000; 2000MO-US04414.  
 PR 24-FEB-2000; 2000MO-US04914.  
 PR 24-FEB-2000; 2000MO-US05004.  
 PR 01-MAR-2000; 2000MO-US05601.  
 PR 20-MAR-2000; 2000MO-US07377.  
 PR 21-MAR-2000; 2000MO-US07532.  
 PR 30-MAR-2000; 2000MO-US08439.  
 PR 17-MAY-2000; 2000MO-US13705.  
 PR 23-MAY-2000; 2000MO-US14042.  
 PR 30-MAY-2000; 2000MO-US14941.  
 PR 02-JUN-2000; 2000MO-US15284.  
 PR 10-NOV-2000; 2000MO-US30873.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Bersani M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 DR P-PSDB; AAU12258.  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX  
 PS Claim 3; Fig 173; 813pp; English.  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 CC  
 CC Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;

Alignment Scores:  
 Pred. No.: 6.35e-94 Length: 1210  
 Score: 848.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-10-082-502-20 (1-162) x AAS21330 (1-1210)  
 QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 DB 603 CGGAGGAGCCAGGATCTCCAGTGTGAGCATGCGGCTCTGGTGGATGAACTGGAATGG 662  
 QY 21 GluIleAlaGluValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
 DB 663 GAAATTCGCCAGGTGGAGCCCAAGAAAGACATTCAGATGGGATCTTCCGATCAATCCA 722  
 QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 DB 723 GATGGCAGCCAGTCAGTGTGTGAGTGCCTTATCCCGCTCAGAGGCCACCTCAGAGAG 782  
 QY 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGluIleAspProSerThr 80  
 DB 783 CTGCTGGAGGAGATATGTGACCGCATGAGAGATGGGAAACGATTCCTTCCACC 842  
 QY 81 HisArgLysAsnTyrValArgValAlaGlyArgAsnGlyLysSerSerGluLeuAspLeu 100  
 DB 843 CATCGCAAGAACTACGATACGTAGTGGGCCGGAATGAGAACTCAGTGAACCTGGACCTA 902  
 QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 903 CAAGCATCCGAATCGATCGATGATTTAGCGCACCTTCAGTGGTGGAGAGCATTT 962  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 DB 963 GTGGAGGAAATACGAGGATGACATTCATTCATTTCCCGAGAGGCTGACAAAGTTTAAA 1022  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
 DB 1023 GACAAACTTTCAGTAAAGCGAAGACAGATCTTGTGACCATGCCCTGCACATATCCATGAT 1082  
 QY 161 GluLeu 162  
 DB 1083 GAGCTA 1088  
 RESULT 10  
 AAX97884  
 ID AAX97884 standard; cDNA: 714 BP.  
 AC AAX97884;  
 XX  
 DT 23-SEP-1999 (first entry)  
 XX  
 DE Human secreted protein encoding cDNA #72.  
 XX  
 KW Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;  
 KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
 OS Homo sapiens.  
 XX  
 PN W09925825-A2.  
 XX  
 PD 27-MAY-1999.  
 XX  
 PF 13-NOV-1998; 98MO-IB01662.  
 XX  
 PR 04-SEP-1998; 98US-0099273.  
 PR 13-NOV-1997; 97US-0066677.  
 PR 17-DEC-1997; 97US-0069957.  
 PR 09-FEB-1998; 98US-0074121.  
 PR 13-APR-1998; 98US-0081563.  
 PR 10-AUG-1998; 98US-0096116.

```

XX (GEST ) GENSET.
XX PA Bouqueleret L, Duclert A, Dumas Milne Edwards J;
XX PI WPI: 1999-347472/29.
XX DR P-PSDB: AAX36200.
XX DR
XX Extended cDNAs encoding secreted proteins
XX PS
XX Claim 1; Page 272; 307pp; English.
XX PS
XX AA97913-97906 represent extended cDNA's which encode novel human
XX secreted proteins (see AAX36129-93622) and which have cytosolic,
XX thrombotic and osteopathic activity. The extended cDNAs can be used to
XX express secreted proteins or parts of them or to obtain antibodies
XX capable of binding to the secreted proteins. They may also be used in
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Uses also include design of expression vectors and secretion vectors.
XX
XX Sequence 714 BP; 203 A; 158 C; 202 G; 151 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 2 26e-93 Length: 714
XX Score: 841.00 Matches: 161
XX Percent Similarity: 99.38% Conservative: 0
XX Best Local Similarity: 99.38% Mismatches: 1
XX Query Match: 99.17% Indels: 0
XX DB: Gaps: 0
XX
US-10-082-502-20 (1-162) x AAX97884 (1-714)
XX
XX 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
XX 93 CGAGAGAGCCAGAGATCTCCACTGTGAGATGAGGCTGTGAGATGAGTACAGATGAG 152
XX
XX 21 GluIleAlaGlnValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
XX 153 GAAATTTGCCAGAGTGGAGCCCAAGAACCAATTCAGATGGGATCTTTCCGATCAATCCA 212
XX
XX 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
XX 213 GATGCGACCACTAGTGTGTGAGAGTGCCTTATGCCCGCTGAGAGCCCACTCCACAGAG 272
XX
XX 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
XX 273 CTGCTGGAGAGATATGTGACCCGATGAAGAGATGGGGAACAGATTGATCTCCACACC 332
XX
XX 81 HisArgLysAsnTyrValArgValValGlyArgGlnGlyLysSerGluLeuAspLeu 100
XX 333 CATCCCAAGAACTAGTACGTACGTACGTGCGCCGCAATGGAGAAATCCAGTCAACTGGACCTA 392
XX
XX 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
XX 393 CAAAGCATCCGAAATGACTCAATATTACCGGACCCCTCAATTTGGCGTGTGGAGCAAT 452
XX
XX 121 ValGlnIuTyrGluAspGluLeuIleGluPhePheSerArgIleAlaAspAsnValLys 140
XX 453 GTGAGAGAAATACGAGAGATCAACTCATTTGAAATCTTTCCCGAGAGCGTGAATGTTAA 512
XX
XX 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160
XX 513 GACAAACTTTTGACGTAAAGGACAGATCTTTGTGACCAATGCCCTGCACATATGCAATGAT 572
XX
XX 161 GluLeu 162
XX 573 GAGCTA 578
XX
XX RESULT 11
XX ID AAX06970 standard; cDNA; 1069 BP.
XX AC AAX06970;

```

```

XX 10-MAY-1999 (first entry)
XX Mouse secretory peptide-9 (Zs1g9) orthologue cDNA.
XX DE
XX Secretory peptide-9; Zs1g9; orthologue; mouse; tumour marker;
XX cancer; therapy; diagnosis; growth enhancer; ss.
XX KM
XX Mus sp.
XX OS
XX Key
XX CDS Location/Qualifiers
XX FT 358..906
XX FT /*tag= a
XX FT sig_peptide 358..417
XX FT /*tag= b
XX FT mat_peptide 418..903
XX FT /*tag= c
XX W09901554-91.
XX
XX 14-JAN-1999.
XX
XX 02-JUL-1998; 98WO-US13859.
XX
XX 17-JUN-1998; 98US-0099005.
XX 03-JUL-1997; 97US-0051704.
XX 03-JUL-1997; 97US-0088088.
XX 19-MAY-1998; 98US-0081338.
XX 19-MAY-1998; 98US-0085983.
XX 17-JUN-1998; 98US-0088899.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX Jaspers SR, Jelinek LJ, Shepard PO, Whitmore TE;
XX WPI: 1999-106055/09.
XX P-PSDB: AAM88476.
XX
XX New mammalian secretory peptide-9 (Zs1g9) - used as a growth
XX enhancer for placenta, liver and heart, and as an indicator of
XX cancer
XX
XX Claim 2; Page 75-77; 85pp; English.
XX
XX This cDNA clone encodes novel mouse secretory peptide-9, or Zs1g9
XX (see AAM88476), an orthologue of novel human Zs1g9 (see AAM88469).
XX Human Zs1g9 is overexpressed in a number of tumours including
XX brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid
XX and lymphoma tumors, and thus can be used as an indicator for
XX cancer. The invention provides polynucleotides (see AAX06968-70)
XX encoding Zs1g9 polypeptides (see AAM88469-77) including mature
XX polypeptides, other processed forms, variants and the mouse
XX orthologues. The Zs1g9 gene, or probes derived from it, can be
XX used to determine if Zs1g9 is present on chromosome 10, and if a
XX mutation has occurred. Antibodies raised against Zs1g9 can be
XX used as diagnostic agents to determine the presence of Zs1g9, and
XX thus the presence of cancer. They can also be labelled with
XX radioisotopes or fused with toxins and used to treat tumors
XX which overexpress Zs1g9. Antisense nucleotides derived from Zs1g9
XX cDNA can also be used to inhibit the growth of tumor cells. Zs1g9
XX proteins can be used to enhance the growth or development of the
XX placenta, heart or liver.
XX
XX Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 8 56e-92 Length: 1069
XX Score: 830.00 Matches: 158
XX Percent Similarity: 98.77% Conservative: 2
XX Best Local Similarity: 97.53% Mismatches: 2
XX Query Match: 97.88% Indels: 0
XX DB: Gaps: 0

```



OY		1	ArgatgSergInAspleuhiscyscgliaIacysAsgAghAlaleuValAsptGlueGluTTP	20
Dd		418	CGAAAGACCAAGAATCTTACACTGTGGAGCTTCACAGGCCTGTGGATGAATTAGAGTG	477
OY		21	GluIlealagInValAsprobslystrHrilegImetGlsySerPhearArgIleasPro	40
Dd		478	GAATTTGCCCGGTGGACCACCAAGAACCATTTCCAGATGGATCCTCCGAATCAATCCA	537
OY		41	AspGlySergInSerVaIvalGluValProTYRAlaArgSergLualahstLeuthrgLu	60
Dd		538	GATGCGACGACCATCTAGTTGTGGAGGTACTTAATGCCCGCTCAGAGGCCACACTCACAGAG	597
OY		61	IeuleuglugluIlecysapArpmetylsglutryGlygluglnIleaspProserThr	80
Dd		558	TTCCTTGAGAGAGGTGTGTACCGAATGAAGAGATACGGGAAACGATGTGACCTTCTACC	657
OY		81	HIsArgLyAsnTrYValarGaValalgaYAgsanGlygluserSergLueAspleu	100
Dd		658	CACCCCAAGAACTACGACTACGCGTCGAGCCGGAAATGGAAATCCAGTAGAAGACTTA	717
OY		101	GinglyIleargIleasPerspIleserGlythrIleuIlysPhehIacysglusertlle	120
Dd		718	CAGGCACTCCGAATTGACACATATACACGGCACCCTCAAGTTTCGtGTGAGACATT	777
OY		121	VAlGIUGlUTryGLuaSpLUleuIllegIuphePheSerArgGluAlaAspaValIlys	140
Dd		778	GTGGAAGAATACAGAGATAGCTTATCGAATTCCTTCACAGAGGCTGCACAACGTATAA	837
OY		141	AspLySLeucysSerIlyArghrAspLeucysAspHisAlaleuHisIleserHisasp	160
Dd		838	GACAAACCTTTGGAGTAAGCGGACAGATCTATGTGACCACAGCCCTGCACAGATCTCACGAT	897
OY		161	Glueu 162	
Dd		898	GAGCTA 903	
RESULT 12				
ID	AAZ08294			
ID	AAZ08294	standard; DNA; 1069 BP.		
AC	AAZ08294;			
DT	07-FEB-2000	(first entry)		
DE		Mouse ortholog gene encoding Zsig9 secretory protein.		
KW		Secretory protein-9; mouse Zsig9; ortholog; overexpression; antagonist;		
KW		antibody; antisense nucleotide; tumour; treatment; receptor;		
KW		radio-label; polypeptide toxin; down-regulation; diagnostic;		
KW		therapeutic; probe; cancer; brain; liver; detection; stomach;		
KW		Lymphoma; ds.		
DS	Mus musculus.			
FH	Key	Location/Qualifiers		
FT	CDS	358..906		
FT		/tag= a		
FT		/product= "Mouse Zsig9 secretory protein"		
FT		/note= "Overexpressed in tumours"		
FT	sig_peptide	358..417		
FT		/tag= b		
FT	mat_peptide	418..903		
FT		/tag= c		
FT		/label= Mature_Zsig9_protein_variant-4		
XX	WO9960405-A1.			
XX				
XX	25-NOV-1999.			
XX				
PF	19-MAY-1999;	99WO-US11107.		
XX				

PR	19-MAY-1998:	9805-0081183.
XX		
PA	(ZYMO ) ZYMOGENETICS INC.	
XX		
PI	Moore EE, Taft DW;	
XX		
DR	WPI: 2000-039447/03.	
XX		
DR	P-PSDB; AAY15136.	
XX		
PT	Detecting tumors using antibodies, antagonists and antisense	
XX	nucleotides to secretory protein-9 (Zs19)	
XX		
PS	Disclosure; Page 37-38; 45pp; English.	
XX		
CC	The present DNA sequence is an ortholog encoding the secretory protein-9	
CC	Zs19 derived from mouse. It is overexpressed in tumors. Antagonists,	
CC	antibodies and antisense nucleotides to Zs19 are useful for detecting	
CC	and treating tumours. The antagonist may be an antibody or receptor to	
CC	Zs19 and it may be radio-labelled or fused to a polypeptide toxin. It	
CC	can be used for down regulating the overexpression of Zs19. The gene	
CC	sequence can be used as nucleic acid probes to detect RNA encoding Zs19	
CC	The Zs19 sequence facilitates improved diagnostic and therapeutic	
CC	techniques for detecting and treating cancers, especially of the brain,	
CC	liver, stomach, lymphoma etc., at an early stage.	

Alignment Scores:		8.56e-92	Length:	1069
Pred. No.:	Score:	830.00	Matches:	158
Percent Similarity:		98.77%	Conservative:	2
Best local Similarity:		97.53%	Mismatches:	2
Query Match:		97.88%	Indels:	0
DB:		21	Gaps:	0

US-10-082-502-20 (1-162) x AAZ08234 (1-1069)

QY	1	ArgArSerGlnAspPLeuNHlSCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp	20
DB	418	CGAAGACCCAAAGATCTTACATCTGTGGACCTTGCAGGGCTCTGGTGGATGAATTAGAGTGG	477
QY	21	GluIleAlaGlnValAspProLysLysThrIleGlnMetGlySerPheArgLysPro	40
DB	478	GAATTTGGCCCGTGGAGCCCAAGAAACCATTCAGATGGATGCTCTCCGAATCATTCGA	533
QY	41	AspGlySerGlnSerValValGluValProTryAlaArgSerGluAlaHisLeuThrGlu	60
DB	538	GATGGCACCCAGTCAAGTTGTGGAGCTACTTATGCCCTTAGAGGCCACCTCAACAG	597
QY	61	LeuLeuGlnGluIleCysAspArgMetLysGluTrpGlyGlnGlnIleAspProSerThr	80
DB	598	TTGCTTGAGAGAGGTGTGTGACCGCAATGAAGAGTACGGGGAAACAGATTGACCTTCTTACC	655
QY	81	HisArgLysAsnTryValArgValValGlyArgAsnGlyLysSerSerGluLeuAspLeu	100
DB	658	CACGCGAAGACTACGTACGCGTGTGAGCCGGAAATGAGAAATCCAGTAAGTACACTTA	711
QY	101	GlnGlyIleArgLysAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle	120
DB	718	CAGGCGATCCCAATTTGACTCAGATATCAAGCGGCAACCTTCGAAGTTGGCTGTGAAGACATT	777
QY	121	ValGlnGlyTrpGluAspGluLeuIleGluPhePheSerArgLysAlaAspAsnValLys	140
DB	778	GTGGAGAAGATACGAGAGAGACTTATCGAATCTTCTTCACAGAGCGTCGACAAAGCTTAAA	833
QY	141	AspLysLeuCysSerLysArgTrpHisPheLysAspHisAlaLeuHisIleSerHisAsp	160
DB	838	GACAAACTTTCACATTAAGCGGACAGATCTATGTACCATGCCCTTGCAACAGATCTCACGAT	897
QY	161	GluLeu 162	
DB	898	GAGCTA 903	

RESULT 13  
ABK36007  
ID ABK36007 standard; cDNA: 657 BP.  
XX  
AC ABK36007;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE cDNA sequence #398 encoding novel human secreted protein.  
XX  
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
KW immune deficiency disorder; blood disorder; inflammatory disorder;  
KW infectious disorder; allergic condition; neurodegenerative disorder;  
KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN NC020017289-A2.  
XX  
PD 18-OCT-2001.  
XX  
PE 29-MAR-2001; 2001WO-US10232.  
XX  
PR 06-APR-2000; 2000US-195605P.  
XX  
PA (GENY ) GENETICS INST INC.  
XX  
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
PI Meberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
PI Clark HF, Fechtel K, Howes SH, Resnick RJ, Gulikova K, Graham JR;  
XX  
DR WPI: 2002-179322/23.  
XX  
PT Six hundred and twenty three polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT  
PS  
XX  
PS Claim 1; Page 297; 393pp; English.  
XX  
CC The present invention relates to the isolation of novel cDNA sequences  
CC which encode human secreted proteins. The cDNA sequences have been  
CC derived from a variety of human tissues. The invention also provides  
CC a method for producing proteins from these polynucleotide sequences.  
CC The proteins are useful for identifying compounds that modulate their  
CC activity and production. The sequences of the invention are  
CC useful for treating diseases such as hyperproliferative disorders  
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
CC haemophilia), and tumours. The polynucleotide sequences of the  
CC invention are also useful in gene therapy. ABK35610-ABK36232 represent  
CC the cDNA sequences of the invention that encode for novel human  
CC secreted proteins.  
XX  
SQ Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;

Alignment Scores:  
Pred. No.: 5,93e-92 Length: 657  
Score: 829.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.76% Indels: 0  
DB: 24 Gaps: 0

US-10-082-502-20 (1-162) x ABK36007 (1-657)  
OY 1 ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValAspGluLeuLysTrp 20  
|||||

DB 49 CGGAGAGCCAGCATCTCCACTGTGAGACATCAGAGGCTCTGTGATGACTAGATGG 108  
OY 21 GIUILEALAGLValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
|||||  
DB 109 GAAATTGCCACGCTGGACCCCAAGAACATGACATGGATCTTTCGGATCAATCA 168  
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
|||||  
DB 169 GATTCACCCAGTCAGTGTGTGGAGGTGCCCTATGCCCGCTCAGAGGCCACATCAGAG 228  
OY 61 LeuLeuGluGluIleIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
|||||  
DB 229 CTGCTGGAGGAGATATGTGACCGGATGAGAGGTATGGGAAACAGATGATCTTCCACC 288  
OY 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyLysSerGluLeuAspLeu 100  
|||||  
DB 289 CATGCCAGAACTAGTACTGCTGTACTGGCCCGAATGGAAATCCAGTAACTGGACCTTA 348  
OY 101 GInGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
DB 349 CAAGCCATCCGAAATGAGTACATGATATAGCCGACCCCTCAAGTTGCGTGTGAGACAT 408  
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
|||||  
DB 409 GTGGAGGATACGAGATGAACTCATTTGAAATCTTTCCGAGAGGCTACAAATGTAA 468  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSerHisAsp 160  
|||||  
DB 469 GACAAACTTTCAGATGAGGAAACAGATCTTTGTGACATGCCCTGCACATATGCGATGAT 528  
OY 161 GluLeu 162  
|||||  
DB 529 GAGCTA 534

RESULT 14  
ID AAX97837  
XX  
AC AAX97837;  
XX  
DT 23-SEP-1999 (first entry)  
XX  
DE Human secreted protein encoding cDNA #25.  
XX  
KW Secreted protein; human; cytostatic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9925825-A2.  
XX  
PD 27-MAY-1999.  
XX  
PE 13-NOV-1998; 98WO-IB01862.  
XX  
PR 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
PA (GENST ) GENSET.  
XX  
PI Bouguetleret L, Duclert A, Dumas Mline Edwards J;  
XX  
DR WPI: 1999-347472/29.  
XX  
DR P-PSDB: AAY36153.  
XX  
PT Extended cDNAs encoding secreted proteins  
XX  
PS Example 28; Page 194-195; 307pp; English.  
XX



BASE COUNT 723 a /note="cloning vector: pHE18SFL13"  
ORIGIN 432 c 477 g 746 t

## Alignment Scores:

Pred. No.:	9,17e-16	Length:	2378
Score:	248.00	Matches:	50
Percent Similarity:	62.30%	Conservative:	26
Best Local Similarity:	40.98%	Mismatches:	44
Query Match:	29.28%	Indels:	2
DB:	9	Gaps:	1

US-10-082-502-21 (1-162) x AK094445 (1-2378)

```
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||||
    :::::|||||
DB 80 GACGGACTGTACTACAGGAGCGATCCCTGCTCAGTCGAGGCGCTTCCTACGGAT 139
    |||||
QY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGluGlnIleAspProSerThr 80
    |||||
    :::::|||||
DB 140 CTTTGGAGAAAGTCTGTGTGCGAATGAAACGACTACAGCTTGAGGAGACCTGTGACG 199
    |||||
QY 81 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeu 100
    |||||
    :::::|||||
DB 200 AAGGAGAGAACTTTCAGAGATTCGCTCCTAGGAAAGAGACAAAATATACCAAGATTT 259
    |||||
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    :::::
    :::::|||||
DB 260 AAAAATTTGATTTTATCTGTATGCTTACAGACCTTGAAATTTGCTGTGAAACTATA 319
    |||||
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    :::::
    :::::|||||
DB 320 ATAGAAGATATGATGAAATATCTCCTATTCGCCAGGAGACACACTATCTAGCT 379
    |||||
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||||
    :::::|||||
DB 380 GACAACTGTGCGAGTGAATAATCAGATCTGTGTGAACCTTCT-----GCTAATCAACT 433
    |||||
QY 161 GluLeu 162
    |||||
DB 434 GAGCTC 439
```

Search completed: March 31, 2003, 03:44:28  
Job time : 1677.7 secs



44607 46803: contig of 2197 bp in length  
\* 46804 46903: gap of unknown length  
\* 46904 49226: contig of 2323 bp in length  
\* 49227 49326: gap of unknown length  
\* 49327 51433: contig of 2107 bp in length  
\* 51434 51533: gap of unknown length  
\* 51534 53763: contig of 2230 bp in length  
\* 53764 53863: gap of unknown length  
\* 53864 56613: contig of 2750 bp in length  
\* 56614 56713: gap of unknown length  
\* 56714 58773: contig of 2060 bp in length  
\* 58774 58873: gap of unknown length  
\* 58874 61527: contig of 2654 bp in length  
\* 61528 61627: gap of unknown length  
\* 61628 63192: contig of 1565 bp in length  
\* 63193 63292: gap of unknown length  
\* 63293 65362: contig of 2070 bp in length  
\* 65363 65462: gap of unknown length  
\* 65463 67438: contig of 1976 bp in length  
\* 67439 67538: gap of unknown length  
\* 67539 70505: contig of 2967 bp in length  
\* 70506 70605: gap of unknown length  
\* 70606 74524: contig of 3919 bp in length  
\* 74525 74624: gap of unknown length  
\* 74625 76737: contig of 2113 bp in length  
\* 76738 76837: gap of unknown length  
\* 76838 80133: contig of 3296 bp in length  
\* 80134 80233: gap of unknown length  
\* 80234 83122: contig of 2889 bp in length  
\* 83123 83222: gap of unknown length  
\* 83223 86731: contig of 3509 bp in length  
\* 86732 86831: gap of unknown length  
\* 86832 89929: contig of 3098 bp in length  
\* 89930 90029: gap of unknown length  
\* 90030 93107: contig of 3078 bp in length  
\* 93108 93207: gap of unknown length  
\* 93208 97419: contig of 4212 bp in length  
\* 97420 97519: gap of unknown length  
\* 97520 101666: contig of 4147 bp in length.  
Location/Qualifiers  
1. 101666  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-31966"

BASE COUNT 24650 a 22934 c 23569 g 23059 t 7454 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3 42e-32 Length: 101666  
Score: 429.00 Matches: 97  
Percent Similarity: 49.25% Conservative: 1  
Best Local Similarity: 48.74% Mismatches: 2  
Query Match: 50.65% Indels: 99  
Gaps: 1  
DB: 2

US-10-082-502-21 (1-162) x AC109891 (1-101666)

OY 49 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluLysAlaAspArg 68  
|||||  
DB 97572 GTTCCCTATAGCCCGCTCAGAGGCCACCTCAGAGAGTTGCTTAGAGAGATATGTACCCA 97631  
|||||  
OY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 88  
|||||  
DB 97632 ATGAAAGAGATATGGGAAACGATTGACCCCTTACACCGCAAGAACTACGTCGCTGT 97691  
|||||  
OY 89 ValSerArgAsnGlyLysSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 108  
|||||  
DB 97692 GTTGGCCGGAACGAGGATCCAGTGAAGTACGATTAAGGGATATCCGAATCGATTGAGAT 97751  
|||||  
OY 109 IleSerGlyThrLeuLysPhe----- 115  
|||||  
DB 97752 ATCAGTGGGACCCCTCAAGTTTGGCGGTAGTTATGGCTGTGCGGTGAGTAGTACGTTCTTG 97811

OY 115 ----- 115  
DB 97812 GGAGTTCATAGGGAATCTGTGGCAGACAGATTGGAGTGGCGGCGATATGTTGCATG 97871  
OY 115 ----- 115  
DB 97872 CTTGTAAACCCGGGAGTGAAGAGAGAGAGAGTGGTGGAGAGATTGGCATTACTACTACT 97931  
OY 115 ----- 115  
DB 97932 TCTAGGCCAGCTGGAGCGCATATCAATAAACCTGTCTATATAATCCCACTGCCACCCA 97991  
OY 115 ----- 115  
DB 97992 AAACAGGAGAGAGAGTGAAGCGGTGTCTGTGCCCCGGCTGAGAGATCTGGGTCACTTCT 98051  
OY 116 ----- 116  
DB 98052 CTGTCTCTTGTGTGGCGCATGAGAGCATTTGTGAGAGATGAGAGATCACTTATTGA 98111  
OY 130 uPhePheSerArgGluAlaAspAsnValLysAspLysLeuCysSerLysArgThr 148  
|||||  
DB 98112 ATTCTTTCAAGAGAGGCTGACACGTTAAAGCAAACTTTGCAGTAAAGCGGACA 98166

RESULT 13  
AF186113 649 bp mRNA linear PRI 13-JAN-2000  
LOCUS AF186113  
DEFINITION Homo sapiens putative secreted protein ZS1G9 (ZS1G9) mRNA, complete cds  
ACCESSION AF186113  
VERSION AF186113.1 GI:6014631  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 649)  
O'Hara, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
O'Hara, P.  
Homo sapiens putative secreted protein  
Unpublished  
2 (bases 1 to 649)  
Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehner, J. and  
O'Hara, P.  
Direct Submission  
Submitted (14-SEP-1999) Biomedical Informatics, ZymoGenetics,  
Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA  
Location/Qualifiers  
1. 649  
/organism="Homo sapiens"  
/db\_xref="dbEST:R15306"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q15"  
/clone="IMAGE:29675"  
1. 649  
/gene="ZS1G9"  
104. 358  
/gene="ZS1G9"  
/codon\_start=1  
/product="putative secreted protein ZS1G9"  
/protein\_id="AA01431.1"  
/db\_xref="GI:6014632"  
/translation="MKGMGLALLGLLGTAMARRSODLHGACRALVDELEWEIAQ  
VDRKTIQSGSRINPDGSGSVETVTPPRKVAHSGFG"

BASE COUNT 175 a 150 c 171 g 149 t 4 others  
ORIGIN

Alignment Scores:  
Pred. No.: 5.63e-17 Length: 649  
Score: 253.00 Matches: 48  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.96% Mismatches: 0

DEFINITION Rattus norvegicus clone CH230-319E6, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 47 unordered pieces.  
 ACCESSION AC109891  
 VERSION AC109891.3 GI:21738196  
 KEYWORDS HTG; HTGS; PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 101666)  
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., All-oman,F.R., Allen,C.,  
 Alebrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barabara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonin,D.,  
 Bowck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Butney,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davis,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Doutheaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Geo,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gortell,D.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,R., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,  
 Homsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratoch,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Licharge,O., Liu,C., Liu,J., Liu,W., Loulseged,H.,  
 Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Mageshwar,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenwo,S., Ogund,M., Okunodu,G.,  
 Oreguine,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Ryles,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sison,I.,  
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tusey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Umanan,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleceyk,R., Wooden,S., Morley,K.,  
 Wu,C., Wu,X., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 TITLE Unpublished  
 JOURNAL Direct Submission  
 REFERENCE 2 (bases 1 to 101666)  
 AUTHORS Morley,K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 REFERENCE 3 (bases 1 to 101666)  
 AUTHORS Morley,K.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 COMMENT On Jul 12, 2002 this sequence version replaced g1:16860222.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GQKt

Center clone name: CH230-319E6  
 ----- Summary Statistics  
 Sequencing Vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 60229 bases at least Q40  
 Consensus quality: 63959 bases at least Q30  
 Consensus quality: 66730 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 47 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \*  
 \* 1  
 \* 1423: contig of 1423 bp in length  
 \* 1424  
 \* 1523: gap of unknown length  
 \* 1524  
 \* 2972: contig of 1449 bp in length  
 \* 2973  
 \* 3072: gap of unknown length  
 \* 3073  
 \* 4718: contig of 1645 bp in length  
 \* 4719  
 \* 4817: gap of unknown length  
 \* 4818  
 \* 6633: contig of 1815 bp in length  
 \* 6732: gap of unknown length  
 \* 6733  
 \* 8040: contig of 1308 bp in length  
 \* 8041  
 \* 8140: gap of unknown length  
 \* 8141  
 \* 10043: contig of 1903 bp in length  
 \* 10044  
 \* 10143: gap of unknown length  
 \* 10144  
 \* 11237: contig of 1094 bp in length  
 \* 11238  
 \* 11337: gap of unknown length  
 \* 12362: contig of 1025 bp in length  
 \* 12363  
 \* 12462: gap of unknown length  
 \* 12463  
 \* 14100: contig of 1638 bp in length  
 \* 14101  
 \* 14200: gap of unknown length  
 \* 14201  
 \* 15840: contig of 1640 bp in length  
 \* 15841  
 \* 15940: gap of unknown length  
 \* 15941  
 \* 17053: contig of 1113 bp in length  
 \* 17054  
 \* 17153: gap of unknown length  
 \* 17154  
 \* 18836: contig of 1683 bp in length  
 \* 18837  
 \* 18936: gap of unknown length  
 \* 18937  
 \* 20137: contig of 1201 bp in length  
 \* 20138  
 \* 20237: gap of unknown length  
 \* 20239  
 \* 21736: contig of 1499 bp in length  
 \* 21737  
 \* 21836: gap of unknown length  
 \* 21837  
 \* 23619: contig of 1783 bp in length  
 \* 23620  
 \* 23720: gap of unknown length  
 \* 23721  
 \* 25907: gap of unknown length  
 \* 25908  
 \* 26007: contig of 2188 bp in length  
 \* 26008  
 \* 27406: gap of unknown length  
 \* 27407  
 \* 27506: contig of 1399 bp in length  
 \* 27507  
 \* 28837: gap of unknown length  
 \* 28838  
 \* 28937: contig of 1331 bp in length  
 \* 28938  
 \* 30229: gap of unknown length  
 \* 30229  
 \* 30329: contig of 1292 bp in length  
 \* 30330  
 \* 30332: gap of unknown length  
 \* 30333  
 \* 32012: contig of 1683 bp in length  
 \* 32013  
 \* 32112: gap of unknown length  
 \* 32113  
 \* 33198: contig of 1086 bp in length  
 \* 33199  
 \* 33298: gap of unknown length  
 \* 33299  
 \* 35519: contig of 2221 bp in length  
 \* 35520  
 \* 35620: gap of unknown length  
 \* 37677: contig of 2058 bp in length  
 \* 37678  
 \* 37777: gap of unknown length  
 \* 38935: contig of 1159 bp in length  
 \* 38936  
 \* 39036: gap of unknown length  
 \* 39037  
 \* 40273: contig of 1237 bp in length  
 \* 40274  
 \* 40373: gap of unknown length  
 \* 41906: contig of 1533 bp in length  
 \* 41907  
 \* 42006: gap of unknown length  
 \* 42007  
 \* 44506: contig of 2500 bp in length  
 \* 44507  
 \* 44508: gap of unknown length



US-10-082-502-21 (1-162) x AC023500 (1-184762)

QY 49 ValProtyAlaArgSerGluAlaHisLeuThrGluLeuLeuGluValaCysAspArg 68  
|||||  
DB 43486 GTCCCTTATGCCCCGCTCAGAGGCCACCTCAGACAGCTCTGTGAGAGATATGTGACCGG 43545  
QY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValaArgVal 88  
|||||  
DB 43546 ATGACAGAGTATGAGGAGAACATTCATCTCCACCCATCGCAAGAACTACGACCTGTA 43605  
QY 89 ValSerArgAsnGlyLysLeuSerGluLeuAspLeuGlnGlyIleArgLysAspSerArg 108  
|||||  
DB 43606 GTGGGCGGAAATGGAATTCAGTGAATGACCTACAAAGGCATCCGAATCGACTCAGAT 43665  
QY 109 IleSerGlyThrLeuLysPheAla----- 116  
DB 43666 ATTAGCGGACCTCTCAAGTTTGGCGGTGAGCTATGGAATCGGATCGGTGTTGTGGACAT 43725  
QY 116 ----- 116  
DB 43726 TAAGGGTGTGTGGAGAACCAATTAGAAATGAGCTGATATCCTTGTCTCCTCTTTGG 43785  
QY 116 ----- 116  
DB 43786 AGGTGAGGACAGAGGCTCTTATTCCTGTGTACACCATTTCTCCCTGGGTGGCAG 43845  
QY 117 CysGluSerIleValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAla 136  
|||||  
DB 43846 TTGAGAGAGCATTTGTGAGGAATGAGAGATGAGATCATTTGAATTTCTTCCGAGAGGCT 43905  
QY 137 AspAsnValLysAspLysLeuCysSerLysArgThr 148  
|||||  
DB 43906 GACAAATGTTAAACAAACTTTCAGTAAAGCGAAC 43941

RESULT 11  
AC090489  
LOCUS AC090489 207424 bp DNA linear ROD 26-JAN-2002  
DEFINITION genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence.  
AC090489  
VERSION AC090489  
KEYWORDS GI:18376849  
HTG.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS 1 (bases 1 to 207424)  
McCormble, W.R., de la Bastide, M., Spiegel, L., Preston, R.,  
Kirchoff, K., Kuhl, K., Nascimben, L., Zutter, T., Ballja, V.,  
Bell, M., Baker, J., Santos, L., Miller, B., Katzenberger, F.,  
Muller, S., King, L., Yang, C., Palmer, L., O'Shaughnessy, A. and  
Dedila, N.  
TITLE Genomic sequence for Mus musculus, clone RP23-104010, complete  
sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 207424)  
AUTHORS McCormble, W.R.  
TITLE Direct Submission  
JOURNAL Submitted (24-FEB-2001) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
3 (bases 1 to 207424)  
McCormble, W.R.  
TITLE Direct Submission  
JOURNAL Submitted (26-JAN-2002) Lita Annenberg Hazen Genome Sequencing  
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA  
On Jan 26, 2002 this sequence version replaced gi:18201765.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest.

FEATURES  
source  
1..207424  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-104010"  
/clone\_11b="RP23-104010"  
/clone\_11c="RP23-104010"  
misc\_feature  
48132..48165  
/note="We believe the assembly to be correct. The sequence  
is a mononucleotide (T) repeat in which the exact number  
of Ts is unknown. Other subclones in the region show one  
fewer T than that represented in the assembly."

BASE COUNT 52613 a 48671 c 49663 g 56477 t

ORIGIN

Alignment Scores:  
Pred. No.: 4,22e-34 Length: 207424  
Score: 451.50 Matches: 100  
Percent Similarity: 49.26% Conservative: 0  
Best Local Similarity: 49.26% Mismatches: 0  
Query Match: 53.31% Indels: 103  
DB: Gaps: 1

US-10-082-502-21 (1-162) x AC090489 (1-207424)

QY 49 ValProtyAlaArgSerGluAlaHisLeuThrGluLeuLeuGluValaCysAspArg 68  
|||||  
DB 75789 GTACCTTATGCCCCGCTCAGAGGCCACCTCAGACAGCTCTGTGAGAGGTGTGTGCCCA 75848  
QY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValaArgVal 88  
|||||  
DB 75849 ATGACAGAGTACGAGGAGAACATTCATCTCCACCCACCGCAAGAACTACGACCGCTC 75908  
QY 89 ValSerArgAsnGlyLysLeuSerGluLeuAspLeuGlnGlyIleArgLysAspSerArg 108  
|||||  
DB 75909 GTAGCGGAAATGGAATTCAGTGAATGACCTTACACCGCAAGAACTACGACCGCTC 75968  
QY 109 IleSerGlyThrLeuLysPheAla----- 116  
DB 75969 ATCAGCGGACCTCTCAAGTTTGGCGGTGAGTACCCTGTGCGCTGAGTACGTCTTCTTG 76028  
QY 116 ----- 116  
DB 76029 GCAGTTATAGGATTCCTCGGACAGAGATGAGATGATGCGAAGTGTGCTGCTAC 76088  
QY 116 ----- 116  
DB 76089 TTGTAACCTTAAGCAGTCCAGGGGCGAGCGTTGGAGATCAGCTTAAGTTGAGGCTGGGC 76148  
QY 116 ----- 116  
DB 76149 CTAGTTCATAGTATGATTCGAGGCGGACGATATAGTAAAGCCCTGTCTATAAAT 76208  
QY 116 ----- 116  
DB 76209 CCCACCCACCCCAAAACAGTAGAGAGTGAAGCCAGTGTCTTCCACACTTGGGA 76268  
QY 117 ----- 116  
DB 76269 TCTGGGTCAATTCCTGTCTTCTTGTGCTCCGACAGTGTGAGGCAATTCGAAATACGAG 76328  
QY 126 AspGluLeuIleGluPheSerArgGluAlaAspAsnValLysAspLysLeuCysSer 145  
|||||  
DB 76329 GATGAGCTTATGCAATTCCTTCTCCAGAGAGCTGACACACTTAAACAAACTTTCAGAT 76388  
QY 146 LysArgThr 148  
|||||  
DB 76389 AACCGGACA 76397

RESULT 12  
AC109891  
LOCUS AC109891 101666 bp DNA linear HTG 13-JUL-2002

QY	116	-----	116
Db 171727	ANGTAAAGGCAAGAGCGCTCTTATTGCGCTGTGTGATACCCATTTCGCCTTGGGTGGCAG	171668	
QY	117	CysgluSerIleValIglUcIuTyrGlusApGluLeuIleGluPhePheSerArgGluAla	136
Db 171667	TTGTGACAGCATTTGTGGAGAAATACGAGATGTAAATCATTTGAATTCCTTTCCGAGAGGCT	171608	
QY	137	AspAsnValIlyAspIlyLeuCySerIlyAsnIgtThr	148
Db 171607	GACATTCCTTAAAGCAAACTTTCAGTAAAGCAACA	171572	
RESULT 10			
AC023500		184762 bp	DNA
LOCUS			linear HTG 24-AUG-2001
DEFINITION			Human sapiens chromosome 12 clone RP11-193H16, WORKING DRAFT
VERSION			AC023500
KEYWORDS			SEQUENCE, 9 unordered pieces.
SOURCE			HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTROP.
ORGANISM			human.
REFERENCE			1 (bases 1 to 184762)
AUTHORS			Muzny D.M., Adams C., Adio-Ondole, B., Altosman, F.R., Allen C., Alsbrooks, S.L., Amaraltinge, H.C., Are, J.R., Ayale, M., Banks, T., Barbara, J., Benton, J., Blmage, K., Blankenburg, K., Bonini, D., Bouck, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Bunley, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carrott, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Deedrich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Go, J., Garcia, A., Garner, T., Garza, N., Gill, R., Girelli, J., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havila, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Hollway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Hulky, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudh, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C., Kratoch, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Lu, W., Louised, H., Lozando, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., McWhiney, E., McLeod, M.P., Medford, M., Meli, G., Metker, M., Miner, G., Miner, Z., Mitchell, T., Mobbbat, K., Morgan, M., Morris, S., Moser, M., Neel, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwoko, S., Ogun, M., Okunodu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatok, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Teltord, B., Thomas, N., Thomas, S., Umanli, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Wolley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D., Weinstein, G., and Gibbs, R.
TITLE			Direct Submission
JOURNAL			Unpublished
REFERENCE			2 (bases 1 to 184762)
AUTHORS			Worley, K.C.
TITLE			Direct Submission
JOURNAL			Submitted (15-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One

```

REFERENCE
  Baylor Plaza, Houston, TX 77030, USA
  3 (bases 1 to 184762)
AUTHORS
  Morley K.C.
TITLE
  Direct Submission
JOURNAL
  Submitted (24-AUG-2002) Human Genome Sequencing Center, Department
  of Molecular and Human Genetics, Baylor College of Medicine, One
  Baylor Plaza, Houston, TX 77030, USA
  On Aug 4, 2002 this sequence version replaced gi:20944227.
COMMENT
  -----
  Genome Center
  Center: Baylor College of Medicine
  Center code: BCM
  Web site: http://www.hgsc.bcm.tmc.edu/
  Drafting Center Code: BCM
  Contact: hgsc-help@bcm.tmc.edu
  -----
  Project Information
  Center project name: HAFc
  Center clone name: RP11-183H16
  -----
  Summary Statistics
  Sequencing vector: Plasmid;
  Sequencing vector: M13;
  Chemistry: Dye-primer Bodipy: 2% of reads
  Chemistry: Dye-terminator Big Dye: 98% of reads
  Assembly program: Phrap; version 0.990329
  Consensus quality: 183845 bases at least Q40
  Consensus quality: 184849 bases at least Q30
  Consensus quality: 185447 bases at least Q20
  Estimated insert size: 159026; sum-of-contigs estimation
  Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
  -----
  * NOTE: Estimated insert size may differ from sequence length
  * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html)
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 9 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  1
  * 2040: contig of 2040 bp in length
  * 2041 2140: gap of unknown length
  * 2141 4170: contig of 2030 bp in length
  * 4171 4270: gap of unknown length
  * 4271 8284: gap of 4024 bp in length
  * 8285 8394: gap of unknown length
  * 8395 14196: contig of 5802 bp in length
  * 14197 14296: gap of unknown length
  * 14297 33076: contig of 18780 bp in length
  * 33077 33176: gap of unknown length
  * 33177 52295: contig of 19119 bp in length
  * 52296 52395: gap of unknown length
  * 52396 84185: contig of 31790 bp in length
  * 84186 84285: gap of unknown length
  * 84286 120023: contig of 35738 bp in length
  * 120024 120123: gap of unknown length
  * 120124 184762: contig of 66639 bp in length.
  Location/Qualifiers
    1..184762 Homo sapiens
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="12"
      /clone="RP11-183H16"
  -----
  BASE COUNT  47425 a 44751 c 45101 g 46673 t 812 others
  ORIGIN
  Alignment Scores:
  Pred. No.: 2.96e-36 Length: 184762
  Score: 472.00 Matches: 98
  Percent Similarity: 65.13% Conservative: 1
  Best Local Similarity: 64.47% Mismatches: 1
  Query Match: 55.73% Indels: 52
  DB: 2 Gaps: 1

```



JOURNAL REFERENCE AUTHORS TITLE	Direct Submission Unpublished 2 (bases 1 to 155023) Worley, K.C. Direct Submission Submitted (11-MAR-2000)
JOURNAL	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE AUTHORS TITLE	3 (bases 1 to 155023) Worley, K.C. Direct Submission Submitted (09-MAY-2002)
JOURNAL	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On May 3, 2002 this sequence version replaced gi:20335511.

[illegible]

RESULT 7	AX64040	1210 bp	DNA	linear	PAT 16-JUL-2002
LOCUS	AX64040				
DEFINITION	Sequence 173 from Patent WO01040466.				
ACCESSION	AX64040				
VERSION	AX64040.1	GI:21899037			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1	Baker, K.P., Bersciani, M., DeGeorge, L., Desnoyers, L., Filvaroff, E., Gao, W. O., Gerritsen, M. E., Goddard, A., Godowski, P. J., Gurney, A. L., Sherwood, S., Smith, V., Stewart, T. A., Tumas, D., Watanabe, C. K., Wood, W. L. and Zhang, Z.	Secreted and transmembrane polypeptides and nucleic acids encoding same	WO 0140466-A 173 07-JUN-2001;	Genentech Inc. (US)	Location/Qualifiers 1. .1210

```

BASE COUNT      261 a      316 c      379 g      254 t
ORIGIN
Alignment Scores:
Pred. No.:      1.75e-75      Length:      1210
Score:          830.00      Matches:      158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%      Mismatches:  2
Query Match:    97.99%      Indels:      0
DB:             6      Gaps:        0

```

US-10-082-502-21 (1-162) X AX464040 (1-1210)

QY	1	ArgArgSerGlnAspLeuHnIscSslYalAcyAspArgAlaLeuValAspGluLeuGlnIutrr	20
Db	603	CGAGGACCCAGAACTCTCATGTGGACATGCAGAGGCTGTGGTGAATGAATCAATGG	662
QY	21	GluIleAlaArgValAspProbyLeuThrIleGlnMetGlySerPheArgIleAspPro	40
Db	663	GAATTTGCCCAAGTGGACCCCAAGAAACCAATTCAGATGGATCTTTTCCGGATCAATCA	722
QY	41	AspGlySerGlnSerValAlaGluValaProTryalaArgSerGluAlaHnIscLeuThrGln	60
Db	723	GATGCACCACTAGTGGTGGAGGTGCTTATGCCCCGTGAGAGGCCCACTCAAGAG	782
QY	61	LeuLeuGlnGluValAcyAspArgMetGlySglUtryGlyGlnGlnIleAspProSerThr	80
Db	783	CTGCCTGGAGAGATGTGTGACCCGATGTGAAGAGTGTGGGAAACATATGATCTCTTCACC	842

QY	81	HISARGYSANPTTYVALARGVALVASERRGNGSLGUSERSERGLULENAPLEU	100
DB	843	CATCCCAAGAACACTACGTACGTCTACTGTGGCCGGAATGAGAAATCCAGAACCTGACCTTA	902
QY	101	GLNGIYLLEARGPLAESPERASPILESERGTYHRLULEULYSPLHEALCYSGIUSERLLE	120
DB	903	CAGGCCATCCGAATCGACTCAGATTATTAGCGGCACCCCTCAAGTTTGCGTGTGAGAGCAATT	962
QY	121	VALGLUGLUTYRTGLUASPGLULEULEGLUPHERPHERSERARGLUALASPSANVALYS	140
DB	963	GTGGGAGGAATACGAGGAAGAACCTCTTAATTCCTTTTCCGAGAGCGTGACAAATGTTTAAA	1022
QY	141	ASPLYSLEUCYSESELYSARGTHRASPLEUCYSAASPISALALEUHLIARGSERHISASP	160
DB	1023	GACAAACTTTTGAGTAACGAAACAGATCTTTGTGCACCATGCCCTGCACATATCGCATAT	1082
QY	161	GLULEU	162
DB	1083	GAGCTTA	1088

RESULT	8
AC025574/c	
LOCUS	155023 bp DNA linear HTG_09-MAY-2002
DEFINITION	Homo sapiens chromosome 12 clone RP11-348M3, WORKING DRAFT SEQUENCE, 12 unordered pieces.
ACCESSION	AC025574
VERSION	AC025574.13 GI:20428723
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens

REFERENCE  
AUTHORS  
Muzny, D.M., Adams, C.C., Adio-Oduola, B., Ali-Usman, F.R., Allen, C.,  
1 (bases 1 to 155023)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

Barbata, J., Benton, J., Blmage, K., Blankenbaur, K., Bonni, D.,  
Bouck, J., Bowle, S., Brivela, M., Brown, E., Brown, M., Bryant, N. P.,  
Bunay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C.,  
Carroll, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R.,  
Daylla, M. L., Davis, C., Day-Carroll, L., Dederich, D. A.,  
Delaney, K. R., Delgado, O., Denn, A. L., Ding, X., Dind, H. H.,  
Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J.,  
Eamhardt, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hoggins, M., Holloway, C., Hollins, B.,  
Homsl, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jackson, L. E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolyvet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kriblovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisleged, H.,  
Lorado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., McWhiney, E., McLeod, M. P., Meador, M., Mei, G., Metker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Morgan, N., Nickerson, E., Nwokkenko, S., Ogul, M., Okunoda, G.,  
Ordunye, N., Oviedo, R., Pace, A., Payton, B., Peary, J., Perez, L.,  
Peterson, L., Pickens, R., Primus, E., Pu, L. L., Quties, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojudoacan, I., Rolfe, M., Ruiz, S., Savery, G.,  
Scheier, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Sytek, A., Tabor, P., Tameis, A., Tameis, K., Tang, H.,  
Tansey, J., Tasey, C., Taylor, T., Taylor, B., Thomas, N., Thomas, S.,  
Umanji, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, C.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S.,  
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y. F., Zhou, Y., Zorrilla, S., Nelson, D.,  
Weinstock, G., and Gdbbs, R.

REFERENCE 1  
AUTHORS Yokoyama-Kobayashi, M., Yamaguchi, T., Sekine, S. and Kato, S.  
TITLE Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank  
JOURNAL Gene 228 (1-2), 161-167 (1999)  
MEDLINE 99173880  
REFERENCE 2 (bases 1 to 814)  
AUTHORS Kato, S.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUN-1998) Seishi Kato, Research Institute of National Rehabilitation Center for the Disabled, Department of Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan (E-mail: seishi@rehab.go.jp, Tel:042-995-3100(ex.2568), Fax:042-995-3132)  
FEATURES  
SOURCE location/Qualifiers  
1. 814  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HP10390"  
/issue\_type="gastric adenocarcinoma"  
/clone\_id="PKA1-meta-1"  
145..693  
/codon\_start=1  
/product="type II membrane protein"  
/protein\_id="BAAF6498.1"  
/db\_xref="GI:4586840"  
/translation="MKMGWLALIGALGTANARRSODLHCACRALVDELEWIAQ VDKRTTOMGSFRINPDGSSQVVEVPAREAHITELLEICDMKRYGQIDPSTHR KNYVVRNGRSESELDLQIRIDSDISGLTKRACSEIVEEDELIEFSSREADNVK DKLSKRTDLCIDNALHSHDEL"  
BASE COUNT 210 a 194 c 241 g 169 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.07e-75 Length: 814  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
Gaps: 0  
US-10-082-502-21 (1-162) x AB015631 (1-814)  
OY 1 ArgtrgserglnaapleuhisCysglYalAcYarGalaleuValaAspGluLeuGluTrp 20  
DB 205 CGAGAGACCGAGATCTCCACTGTGAGACAGAGGCGTCTGTGAGATGAGAAATGG 264  
OY 21 GlutleAlaargValaAspProlyLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 265 GAAATTCGCCAGGTGAGACCCCAAGAGACCATTCAGATGAGATCTTCCGGATCAATCCA 324  
OY 41 AspGlySerGlnSerValIleGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 325 GATGGACGACGATGAGTGGTGGAGTGCCTTATGCCGCCCTCAGAGGCCCACTCAGAG 384  
OY 61 LeuleuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 80  
DB 385 CTGCTGAGAGATATGATGACCGGATGAGAGATGAGGGAACAGATTGATCTTCCACC 444  
OY 81 HisArgLysAsnIlyrValArgValIleSerArgAsnGlyGlySerGluLeuAspLeu 100  
DB 445 CATGGCAAGAACTGATGACGTGTAGTGGCCGGAATGAGATCCAGTAAGTGGACCTA 504  
OY 101 GlnGlyIleArgGlyLeaSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 505 CAAAGCATCCCAATCGACTCGATATTAAGCGCACCTCAAGTTGGCTGTGAGAGCAATT 564  
OY 121 ValGluGluTrpGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 565 GTGAGAGAAATGACGAGATGAACTCATTTGAATCTTTCCCGAGAGCGTGAACAATGTTAA 624  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisLysSerHisAsp 160

DB 625 GACAAACTTTCAGTACGACGAGACAGATCTTGTGACCATGCCCTGCACATATCCGATGAT 684  
OY 161 GlutLeu 162  
DB 685 GAGCTA 690  
RESULT 6  
AY032624  
LOCUS AY032624 1080 bp mRNA linear PRI 18-APR-2002  
DEFINITION Homo sapiens saposin-like protein mRNA, complete cds.  
ACCESSION AY032624  
VERSION AY032624.1 GI:20196198  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1080)  
AUTHORS Bornhauser, B.C., Olsson, P.-A. and Lindholm, D.  
TITLE NSAP is a novel saposin-like protein that interacts with MTR and stimulates neurite outgrowth  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1080)  
AUTHORS Olsson, P.-A. and Lindholm, D.  
TITLE Direct Submission  
JOURNAL Submitted (13-APR-2001) Neuroscience, Uppsala University, Husargatan 3, Uppsala 75123, Sweden  
FEATURES  
SOURCE location/Qualifiers  
1. 1080  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q15"  
419..967  
/note="NSAP; encodes a type 2 membrane signal and a putative ER retention motif"  
/codon\_start=1  
/product="saposin-like protein"  
/protein\_id="AAK38148.1"  
/db\_xref="GI:20196198"  
/translation="MKMGWLALIGALGTANARRSODLHCACRALVDELEWIAQ VDKRTTOMGSFRINPDGSSQVVEVPAREAHITELLEICDMKRYGQIDPSTHR KNYVVRNGRSESELDLQIRIDSDISGLTKRACSEIVEEDELIEFSSREADNVK DKLSKRTDLCIDNALHSHDEL"  
BASE COUNT 238 a 269 c 343 g 230 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.52e-75 Length: 1080  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
Gaps: 0  
US-10-082-502-21 (1-162) x AY032624 (1-1080)  
OY 1 ArgtrgserglnaapleuhisCysglYalAcYarGalaleuValaAspGluLeuGluTrp 20  
DB 479 CGAGAGACCGAGATCTCCACTGTGAGACAGAGGCGTCTGTGAGATGAGAAATGG 538  
OY 21 GlutleAlaargValaAspProlyLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 539 GAAATTCGCCAGGTGAGACCCCAAGAGACCATTCAGATGAGATCTTCCGGATCAATCCA 598  
OY 41 AspGlySerGlnSerValIleGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 599 GATGGACGACGATGAGTGGTGGAGTGCCTTATGCCGCCCTCAGAGGCCCACTCAGAG 658  
OY 61 LeuleuGluGluValCysAspArgMetLysGluTrpGlyGluGlnIleAspProSerThr 80  
DB 659 CTGCTGAGAGATATGATGACCGGATGAGAGATGAGGGAACAGATTGATCTTCCACC 718



\* 150247 150346: gap of unknown length  
\* 150347 153410: contig of 3064 bp in length  
\* 153411 153510: gap of unknown length  
\* 153511 155739: contig of 2229 bp in length  
\* 155740 155839: gap of unknown length  
\* 155840 158198: contig of 2359 bp in length.

FEATURES  
source  
1. 158198  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP11-764L14"

BASE COUNT 46057 a 30922 c 31747 g 49030 t 442 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1,39e-74 Length: 158198  
Score: 847.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-082-502-21 (1-162) x AC012013 (1-158198)

OY 1 ArgArgSerGlnAspLeuH1scysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
|||||  
DB 152171 CGAAGAGCCAGCATCTACACTGTGAGGCTGAGGCGCTGTGTGATGATTAAGAGTGG 152230  
|||||  
OY 21 GlllAlaIArgValAspProlySlyThrIleGlnMetGlySerPheArgIleAspPro 40  
|||||  
DB 152231 GAATATGCCCGGTGGACCCCAAGAACCATTCACATGGGATCTCCGATCAATCCA 152290  
|||||  
OY 41 AspGlySerGlnSerValValGluValProTyAlaArgSerGlnAlaHisLeuThrGlu 60  
|||||  
DB 152291 GATGCAAGCCAGCATCTGTGGAGTACCTTATGCCCGCTCAGAGGCCCACTCAGAGAG 152350  
|||||  
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyGlyGluGlnIleAspProSerThr 80  
|||||  
DB 152351 TTGCTTGAGAGAGTGTGTGACCAATGAAGAGATGACGGGAGACAGATGACCTTCTAC 152410  
|||||  
OY 81 HisArgLysAsnTyrrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
|||||  
DB 152411 CACCGCAAGAACTACGTACGCGCTGTGACCGGAATGAGAAATCCAGTAACTAGACTTA 152470  
|||||  
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
DB 152471 CAGGCGATCCGAATTCACATCAGATATCAGCGGACCCCTCAAGTTTGGCTGTGAGAGCATT 152530  
|||||  
OY 121 ValGluGluTyrrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
|||||  
DB 152531 GTGCAAGAAATACAGAGATGAGCTTATCGAATCTCTCCAGAGAGGCTGACAACTTAA 152590  
|||||  
OY 141 AspLysLeuLysSerIleAspArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
|||||  
DB 152591 GACAAACTTTGAGTAAAGGAGACAGATCTATGAGACCATGCCCTGCACAGATCTCAGAT 152650  
|||||  
OY 161 GluLeu 162  
|||||  
DB 152651 GAGCTA 152656

RESULT 4  
AX440456 814 bp DNA linear PAT 28-JUN-2002  
LOCUS AX440456 Sequence 309 from Patent WO0190154.  
DEFINITION AX440456  
ACCESSION AX440456  
VERSION AX440456.1 GI:21665266  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1

AUTHORS Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secrist, H.,  
Lodes, M.J., Algate, P.A., Fling, S.P., Mannion, J., Benson, D.R. and  
Carter, D.  
TITLE Compositions and methods for the therapy and diagnosis of ovarian  
cancer  
JOURNAL Patent: WO 0190154-A 309 29-NOV-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source  
1. 814  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 210 a 194 c 241 g 169 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,07e-75 Length: 814  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: 6 Gaps: 0

US-10-082-502-21 (1-162) x AX440456 (1-814)

OY 1 ArgArgSerGlnAspLeuH1scysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
|||||  
DB 205 CGAAGAGCCAGCATCTCCACTGTGAGCATGAGGCGCTGTGTGATGATTAAGAGTGG 264  
|||||  
OY 21 GlllAlaIArgValAspProlySlyThrIleGlnMetGlySerPheArgIleAspPro 40  
|||||  
DB 265 GAATATGCCCGGTGGACCCCAAGAACCATTCACATGGGATCTCCGATCAATCCA 324  
|||||  
OY 41 AspGlySerGlnSerValValGluValProTyAlaArgSerGlnAlaHisLeuThrGlu 60  
|||||  
DB 325 GATGCAAGCCAGCATCTGTGGAGTACCTTATGCCCGCTCAGAGGCCCACTCAGAGAG 384  
|||||  
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyGlyGluGlnIleAspProSerThr 80  
|||||  
DB 385 CTGCTGAGAGAGATATGTGACCCGATGAAGAGATGAGGAGAACAGATGATCTTCCACC 444  
|||||  
OY 81 HisArgLysAsnTyrrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
|||||  
DB 445 CATGCAAGAACTACGTACGCTGTGAGGCGGAAATGAGAAATCCAGTAACTAGACTTA 504  
|||||  
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
DB 505 CAGGCGATCCGAATTCACATCAGATATCAGCGGACCCCTCAAGTTTGGCTGTGAGAGCATT 564  
|||||  
OY 121 ValGluGluTyrrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
|||||  
DB 565 GTGCAAGAAATACAGAGATGAGCTTATCGAATCTCTCCAGAGAGGCTGACAACTTAA 624  
|||||  
OY 141 AspLysLeuLysSerIleAspArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
|||||  
DB 625 GACAAACTTTGAGTAAAGGAGACAGATCTTATGAGACCATGCCCTGCACAGATCTCAGAT 684  
|||||  
OY 161 GluLeu 162  
|||||  
DB 685 GAGCTA 690

RESULT 5  
AB015631 814 bp mRNA linear PRI 26-JUL-2001  
LOCUS AB015631 Homo sapiens mRNA for type II membrane protein, complete cds,  
DEFINITION clone:HP10390.  
ACCESSION AB015631  
VERSION AB015631.1 GI:4586839  
KEYWORDS type II membrane protein.  
SOURCE Homo sapiens gastric adenocarcinoma CDNA to mRNA,  
clone.lib:pkal-meta-1 clone:HP10390.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.





LOCUS	BC008261	798 bp	mRNA	Linear	ROD 07-AUG-2002
DEFINITION	Mus musculus, transmembrane protein 4, clone MGC:6853				
ACCESSION	IMAGE:2650612, mRNA, complete cds.				
VERSION	BC008261				
KEYWORDS	BC008261.1 GI:14198400				
SOURCE	MGC.				
ORGANISM	house mouse.				
REFERENCE	Mus musculus				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
TITLE	1 (bases 1 to 798)				
JOURNAL	Straussberg, R.				
REMARK	Direct Submission				
COMMENT	Submitted (22-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
	NIH-MGC project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>				
	Contact: MGC help desk				
	Email: <a href="mailto:cgapbs-rt@mail.nih.gov">cgapbs-rt@mail.nih.gov</a>				
	Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center				
	Center code: BCM-HGSC				
	Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a>				
	Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a>				
	Gunnarstne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M., Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.				
FEATURES	clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>				
SOURCE	Series: IRAC Plate: 5 Row: b Column: 23				
	This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9903606.				
	Location/Qualifiers				
	1..798				
	/organism="Mus musculus"				
	/db_xref="taxon:10090"				
	/map="129,C57BL/6J,FVB/N"				
	/clone="MGC:6853 IMAGE:2650612"				
	/tissue_type="Mammary tumor. Brca1-/-l; MMTV-Cre model. 10 months old, gross tissue."				
	/clone_idb="NCI CGAP_Mam3"				
	/lab_host="DH10B"				
	/note="Vector: pCMV-SPORT6"				
	115..663				
	/codon_start=1				
	/product="transmembrane protein 4"				
	/protein_id="AA08261.1"				
	/db_xref="GI:14198401"				
	/db_xref="LOCUSID:56530"				
	/translation="MKMGWLLALLGLVLTGAMARQDLCGACRALVDELEMEIARVDPKRTIQMSFRINPDGQSIVVEPYARSPAHITLLEVCDDMKKYGEGIDISTHRRKNYRVYRNGSESELDLOGIRIDSDISGTLKFACESIVEVEDELLIEFSEADNVKWDKLSKRTDLCDHAIHRSDEL"				
CDS					
	115..663				
	/codon_start=1				
	/product="transmembrane protein 4"				
	/protein_id="AA08261.1"				
	/db_xref="GI:14198401"				
	/db_xref="LOCUSID:56530"				
	/translation="MKMGWLLALLGLVLTGAMARQDLCGACRALVDELEMEIARVDPKRTIQMSFRINPDGQSIVVEPYARSPAHITLLEVCDDMKKYGEGIDISTHRRKNYRVYRNGSESELDLOGIRIDSDISGTLKFACESIVEVEDELLIEFSEADNVKWDKLSKRTDLCDHAIHRSDEL"				
BASE COUNT	227 a 182 c 228 g 161 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	1,93e-77	Length:	798		
Score:	847.00	Matches:	162		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	10	Gaps:	0		

Db	175	CGAAGGACCCAAAGTCTACACTGTGGAGCTTGCAGGGCTCTGGTGGATGAATTGAGTGG	234
QY	21	GIUIleaIaArVaIaAspProLyAsThrIIleGIIneIeGlySerPheArgIIeaSPro	40
Db	235	GAATTTGGCCCCGfGGAGACCCCAAGAACCAATTCAGATGGGATCCTTCCGAATCAATCA	294
QY	41	AspGlySerGIInSerVaIaGIuVaIProTfAlaArGSerGIuAlaIhISLeuThrGIu	60
Db	295	GATGGCAGCCAGTCAAGTTGTGGAGTACCTTATGCCCTCAGAGGCCACCTCACAGAG	354
QY	61	LeuGIuGIuVaIcYsAspArgMetLySGIuArgLyGIuGIuIIleAspProSerThr	80
Db	355	TTGCTTGAGGAGGGTGTGGACCGAATGAAGAGTACGGGGGAACAGATTGACCCTTTTACC	414
QY	81	HIaSIrGLySAaNTfYrVaIaArVaIaIaSerArGAnGIyGIuSerSerGIuLeuAspLeu	100
Db	415	CACGGCAAGACTACGTACGCTGCTGAGCCGGAAATGAGAAATCCAGTAACTAGACTTA	474
QY	101	GIInGIyIIeaArgIIeaSProLyAsThrIIleGIIneIeGlySerPheArgIIeaSPro	120
Db	475	CAGGGCATCCCAATTTGACTCAGATATACGGCGACCCCTCAAGTTTGCCTGTGAGAGCAT	534
QY	121	ValGIuGIuArgLyGIuVaIaAspGIuLeuIIleGIuIhPheSerArgGIuAlaAspAsnVal	140
Db	535	GTGGAAGATACGAGAGTACGCTTATCCAAATCTTCTCCAAAGAGCGTGACACGTTTAA	594
QY	141	AspLyLeuCySserLyAsrGThrAspLeuCySAspHISAlaLeuHISArgSerHISAsp	160
Db	595	GACAACTTTCACAGTAAAGCGACAGATCTTATGTACCACTGCCCTGCACAGATCTCACAT	654
QY	161	GIuLeu 162	
Db	655	GAGCTA 660	
RESULT 2	AF186115		
LOCUS	AF186115	1017 bp	linear
DEFINITION	Mus musculus putative secreted protein ZS1G9 (Zs1g9) mRNA, complete cds.		
ACCESSION	AF186115		
VERSION	AF186115.1	GI:6014635	
KEYWORDS	Mus musculus.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 1017)		
AUTHORS	Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehner, J. and O'Hara, P.		
TITLE	Mus musculus putative secreted protein		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1017)		
AUTHORS	Sheppard, P., Jelinek, L., Whitmore, T., Blumberg, H., Lehner, J. and O'Hara, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-SEP-1999) Biomolecular Informatics, ZymoGenetics, Inc., 1201 Eastlake Ave. East, Seattle, WA 98102, USA		
FEATURES	location/Qualifiers		
source	1..1017		
	/organism="Mus musculus"		
	/db_xref="dbEST:AA049839"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:476061"		
	1..1017		
	/gene="Zs1g9"		
	343..891		
	/gene="Zs1g9"		
	/codon_start=1		
	/product="putative secreted protein ZS1G9"		
	/protein_id="AA01433.1"		
	/db_xref="GI:6014636"		
	/translation="MKGWMLALLGLVLTGAVANARRSDLHGACRALVDLEMEIAR		

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:24:04 ; Search time 1535.7 Seconds  
(without alignments)  
3070.029 Million cell updates/sec

Title: US-10-082-502-21  
Perfect score: 847  
Sequence: 1 RRSQDLHCACGACALVDELEW.....LCGRKRDLCDAHRSDEL 162

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ -p2n.model -DEV-xip  
-O/cg12.1/USPTO.spool/US10082502/rnatc.24032003.135059.6547/app.query.fasta.1.1308  
-DB-GenEmbl -OFMT-fastap -SUFFIX-rge -MINMATCH-0.1 -LOOPEXT-0 -LOOPEXT-0  
-UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45  
-DOCALLGN-200 -THR\_SCORE-pct -THR\_MAX-100 -THR\_MIN-0 -ALIGN-15 -MODE-LOCAL  
-OUTFMT-pio -NOR-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000  
-USER-US10082502.qc1n.1.1.4758.qrunat.24032003.135059.6547 -NCP-6 -ICPU-3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES-0 -WAIT -LONGLOG -DEV\_TIMECUT-120  
-WARN\_TIMECUT-30 -THREDS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7  
-YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hgt.\*  
3: gb\_in.\*  
4: gb\_com.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_com.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vl.\*  
30: em\_hgt\_hum.\*  
31: em\_hgt\_inv.\*  
32: em\_hgt\_other.\*  
33: em\_hgt\_mus.\*  
34: em\_hgt\_pln.\*  
35: em\_hgt\_rnd.\*  
36: em\_hgt\_mam.\*  
37: em\_hgt\_vrt.\*  
38: em\_sy.\*  
39: em\_hgtgo\_hum.\*  
40: em\_hgtgo\_mus.\*  
41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	847	100.0	798	10	BC008261 Mus muscu
2	847	100.0	1017	10	AF186115 Mus muscu
3	847	100.0	158198	2	AC012013 Homo sapi
4	830	98.0	814	6	AX440456 Sequence
5	830	98.0	814	9	AB015631 Homo sapi
6	830	98.0	1080	9	AY032624 Homo sapi
7	830	98.0	1210	6	AX464040 Sequence
8	474.5	56.0	155023	2	AC025574 Homo sapi
9	472	55.7	182350	2	AC073896 Homo sapi
10	472	55.7	184762	2	AC023500 Homo sapi
11	451.5	53.3	207424	10	AC090489 Genomic s
12	429	50.6	101666	2	AC109891 Rattus no
13	253	29.9	649	9	AF186113 Homo sapi
14	253	29.9	824	9	BC001027 Homo sapi
15	248	29.3	2378	8	AK094445 Homo sapi
16	228.5	27.0	69753	2	AC103156 Rattus no
17	204	24.1	996	3	AY094698 Drosophi
18	193.5	22.8	40824	3	U13070 Caenorhbd
19	156.5	18.5	622	5	AB055671 Danto rer
20	155	18.3	785	8	AY088187 Arbidops
21	154	18.2	564	6	AX440375 Sequence
22	141.5	16.7	175963	3	AC009379 Drosophi
23	141.5	16.7	279530	3	AE003518 Drosophi
24	140	16.5	55359	2	AC012854 Drosophi
25	140	16.5	172904	3	AC007414 Drosophi
26	140	16.5	275390	3	AE003831 Drosophi
27	138.5	16.4	1286	3	AY060987 Drosophi
28	133	15.7	161566	9	AC008060 Homo sapi
29	132	15.6	744	6	AR142811 Sequence
30	132	15.6	746	6	AX136616 Sequence
31	132	15.6	747	6	BD006701 Novel pol
32	132	15.6	1413	9	BC019903 Homo sapi
33	132	15.6	1419	6	AX136423 Sequence
34	132	15.6	1470	6	AX376412 Sequence
35	132	15.6	1470	6	AX454716 Sequence
36	132	15.6	1470	6	AX491194 Sequence
37	132	15.6	1512	6	AR142810 Sequence
38	132	15.6	1547	9	BC032339 Homo sapi
39	124	14.6	76187	2	AC015219 Drosophi
40	117	13.8	168418	9	AC009403 Homo sapi
41	116	13.7	1260	6	AX380553 Sequence
42	116	13.7	181896	2	AC114126 Rattus no
43	116	13.7	201649	2	AC120813 Rattus no
44	115.5	13.6	201649	2	AC120813 Rattus no
45	113	13.3	210631	2	AC129184 Mus muscu

RESULT 1

## ALIGNMENTS



Db 165 GAAATGCCCGCGTGACCCCAAGACACATTCAGATGGATCCTTCGAATCAATCCA 224  
QY 41 AspglySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
|||||  
Db 225 GATGGCAGCCAGTGTGTGTGAGGTACTTATGCCCTCAGAGGCCACCTCCAGAG 284  
QY 61 LeuleuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
|||||  
Db 285 TTGCTTGAGGAGGTGTGTGACCGAATGAAAGAGTACGGGAA-CAGATTGA-CCTTCTACC 342  
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100  
|||||  
Db 343 CACCGCAAGAACTACTACGCTACGCTGTGAGCCGGAATGGAGATCCAGTGAACCTAGACTTA 402  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
|||||  
Db 403 CAGGGCATCCGAAATTGACTCAGATATCANGGCAACCTCAAGTTGCGTGTGAGAGCAT 462  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
|||||  
Db 463 GTGGAAGAAATACGAGGATGACTTATCGAATCTCTCCAGAGAGGTGACACACTTAA 522  
QY 141 AsplysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSer-HisAs 160  
|||||  
Db 523 CACAACCTTTGCAGTAAACGCGACAGATCTATGTGACCATCCCTGCACAGATCCTCACGA 582  
QY 160 pGluLeu 162  
:|||||  
Db 583 AGAGCTC 589

Search completed: March 30, 2003, 22:43:01  
Job time : 130.151 secs

CC AX97813-1X97906 represent extended cDNAs which encode novel human  
CC secreted proteins (see AY36129-Y36222) and which have cytostatic,  
CC thrombotic and osteoprotectin activity. The extended cDNAs can be used  
CC to express secreted proteins or parts of them or to obtain antibodies  
CC capable of binding to the secreted proteins. They may also be used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC uses also include design of expression vectors and secretion vectors.

**SQ** Sequence 832 BP; 224 A; 194 C; 243 G; 168 T; 3 other;

**Alignment Scores:**

Pred. No.:	4.49e-85	length:	833
Score:	781.00	Matches:	150
Percent Similarity:	96.36%	Conservative:	3
Best local Similarity:	94.55%	Mismatches:	3
Query Match:	92.21%	Indels:	3
DB:	20	Gaps:	0

US-10-082-502-21 (1-162) x AAX97837 (1-832)

QY	1	ArgTsrGsrGlnAspLeuHnIAsCysGlyValAcysArgAlaLeuValAspGluLeuGlnTr	20
Db	208	CGGAGGACCGCAGGATCTCCACTGTGGAGCATGCAAGGGCTCTGTGTGAATCAAGAAATG	267
QY	20	pgLIuIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPr	40
Db	268	GGAATATGGCCAGGTGGACCCCAAGAAAGCAATTCAGATGGGATCTTCCGATCAATCC	327
QY	40	oAbpGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThGl	60
Db	328	AGATGGCAGCCAGTCAGTGTGTGGAGGTGCCCTTATGCCCGCTCAGAGGCCACCTCACAGA	387
QY	60	uLeuLeuGlnGluValAlCysAspArgMetLysGluTyrTyrGlnGlnIleAspProSerTh	80
Db	388	GCTCTCTGAGAGATATATGATGCAGCCGATGAAGAGATATGGGGAAACAGATGATCTTCCAC	447
QY	80	rHisArgLysAsnTyrValArgValSerArgAsnGlyGlnSerSerGluLeuAspLe	100
Db	448	CCATCGCAAGAACTAGCTAGCTGTAGTGGGCCGGAATGGAAATCCAGTAATCGAATCGACT	507
QY	100	uGlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLys-PheAlaCysGlnSerI	120
Db	508	ACAAGGCATCCGAATCGACTCAGATATTGAGCGGCACCCCAAGBPTTGGCTGCGGAACA	567
QY	120	leValGlnGluTyrGlnLysAspGlnLeuIleGlnPhePheSerArgGluAlaAspAsnValL	140
Db	568	TTGTGGAGGAATACGAGAGATGAATCTATTGAATCTTTTCCGAGAGGGCTGACAAATGTTA	627
QY	140	ysaAspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSer-His	159
Db	628	AAGCAAACTTTGCAAGTAAACGAAACAGATCTTTGTATACCATGCCCTGCACAAATGGGAT	687
QY	160	AspGlnLeu 162	
Db	688	GATGAGCTA 696	
RESULT 15			
AAA43525			
XX	AAA43525	standard; cDNA; 592 BP.	
XX	AAA43525;		
XX	21-AUG-2000	(first entry)	
DE	Mouse secreted expressed sequence tag SEQ ID NO:100.		
KW	Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;		
KW	expressed sequence tag; EST; probe; chemotactic; proliferative;		
KW	immunomodulatory; haematopoietic; chemokinetiic; analgesic; haemostatic;		
KW	thrombolytic; antiinflammatory; cyostatic; antibacterial; antifungal;		
KW	antiviral; antidiabetic; antiaesthetic; vulnerary; antiparkinsonian;		
KW	antitumor; osteoparalic; neuroprotective; nootropic; antipsoriatic;		
KW	cerebroprotective; anticonvulsant; antidepressant; gene therapy;		

KM vaccine autoimmune disorder; multiple sclerosis; allergic condition  
KM insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KM lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KM central nervous system disorder; Alzheimer's disease; stroke;  
KM Parkinson's disease; Huntington's disease; coagulation disorder;  
KM hemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KM tumour; infection; depression; psoriasis; ss.

**Mus musculus.**

PN WO200021991-A1.

PD 20-APR-2000

PF 15-OCT-1999; 99MO-US24206.

PR 15-OCT-1998; 98US-0104436.

PA (GEMY ) GENETICS INST INC.

PI Jacobs K, McCoy JM, Lavalie ER, Collins-Racle LA, Evans C;

XX

XX

PT expressed sequence tags (ESTs), useful for treating various disorders

XX

XX

CC · sequence tags (SESTs), isolated from human, mouse, chicken and rat

these sources. The assays can have a range of activities depending on the tissues they were isolated from. The activities include:

- chemotactic; proliferative; immunomodulatory; haematopoietic;
- chemokinetic; anagestic; haemostatic; thrombolytic; antiinflammatory;
- cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
- antiallergic; vulnery; antitumor; osteoprotic; neuroprotective;
- neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;
- anticonvulsant; and antidepressant. The SESs can be used for gene therapy and in vaccines. The SESs are useful as probes for the identification and isolation of full-length cDNAs and genomic DNA molecules which correspond to the SESs. Proteins encoded by the SESs are useful in assays for determining biological activity and raising antibodies. They may be useful for treatment of autoimmune disorders (multiple sclerosis, insulin dependent diabetes), allergic conditions (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers, osteoporosis, osteoarthritis, central nervous system disorders (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation disorders (thrombophilia, thrombosis), inflammatory disorders (Crohn's disease), tumours, bacterial, fungal or viral infections, depression and psoriasis. AAA45926 to AAA45931 represent linker variants which are given in the exemplification of the present invention.

**SQ** Sequence 592 BP; 159 A; 139 C; 167 G; 123 T; 4 other,

**Alignment Scores:**

Pred. No.:	3,77e-03	Length:	59
Score:	763.50	Matches:	15
Percent Similarity:	95.71%	Conservative:	1
Best Local Similarity:	95.09%	Mismatches:	6
Query Match:	90.14%	Indels:	3
DB:	21	Gaps:	0

US-10-082-502-21 (1-162) x AAA43525 (1-592)

OY 1 ArgAdgSerGlnAspLeuHicScysGlnAlaCysArgAlaLeuValAspGlnLeuGlnIleuTrp 20  
Db 105 CGAAGGACCAANATCTACATGTGTGGACCTTGCAAGGCTCTGTGGATGAATTAAGCTGG 16  
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40

RESULT 13  
ABK36007  
ID ABK36007 standard; cDNA: 657 BP.  
XX  
AC ABK36007:  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE cDNA sequence #398 encoding novel human secreted protein.  
XX  
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
KW immune deficiency disorder; blood disorder; inflammatory disorder;  
KW infectious disorder; allergic condition; neurodegenerative disorder;  
KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200177289-A2.  
XX  
PD 18-OCT-2001.  
XX  
PE 29-MAR-2001; 2001WO-US10232.  
XX  
PR 06-APR-2000; 2000US-195605P.  
XX  
PA (GENY ) GENETICS INST INC.  
XX  
PI Jacobs K, McCoy JM, Lavalie ER, Collins-Racie LA, Evans C;  
PI Meberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
PI Clark HF, Reichel K, Howes SH, Resnick RJ, Gultukota K, Graham JR;  
DR WPI: 2002-179322/23.  
XX  
XX Six hundred and twenty three polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT  
XX  
PS Claim 1: Page 297; 393pp; English.  
XX  
XX The present invention relates to the isolation of novel cDNA sequences  
CC which encode human secreted proteins. The cDNA sequences have been  
CC derived from a variety of human tissues. The invention also provides  
CC a method for producing proteins from these polynucleotide sequences.  
CC The proteins are useful for identifying compounds that modulate their  
CC activity and production. The sequences of the invention are  
CC useful for treating diseases such as hyperproliferative disorders  
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined  
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple  
CC sclerosis), blood disorders (e.g. thrombocytopenia), inflammatory  
CC disorders (e.g. arthritis), infectious disorders (e.g. hepatitis),  
CC allergic conditions (e.g. asthma), neurodegenerative disorders (e.g.  
CC Alzheimer's disease), liver fibrosis, coagulation disorders (e.g.  
CC hemophilia), and tumours. The polynucleotide sequences of the  
CC invention are also useful in gene therapy. ABK3510-ABK3622 represent  
CC the cDNA sequences of the invention that encode for novel human  
CC secreted proteins.  
XX  
SO Sequence 657 BP; 177 A; 149 C; 185 G; 146 T; 0 other;

Alignment Scores:  
Pred. No.: 2,49e-89 Length: 657  
Score: 815.00 Matches: 155  
Percent Similarity: 97.53% Conservative: 3  
Best Local Similarity: 95.68% Mismatches: 4  
Query Match: 96.22% Indels: 0  
DB: 24 Gaps: 0

US-10-082-502-21 (1-162) x ABK36007 (1-657)

OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaLeuArgAlaLeuValAspGluLeuGluTTP 20  
|||||

DB 49 CGAGAGACCCGAGATCTCCACTGTGGAGACATGACAGGCTCTGGTGATGAATCGAATGC 108  
OY 21 GluIleAlaArgValAspProLysLeuThrIleGlnMetGlySerPheArgIleAspPro 40  
DB 109 GAATTTGCCACGATGGAGACCCCAAGAGACATGACAGATGGATCTTCCGATCATCA 168  
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 169 GATTGACGCCAGTACGTGTGTGGAGTGTCCCTATGCCCCCTAGAGGCCCATCATCAGAG 228  
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
DB 229 CTGCTGGAGGAGATATGTGACCGGATGAGAGATATGGGAAACAGATTGATCTTCCACC 288  
OY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
DB 289 CATGCAAGAACTACGTACGTAGTGTGGCCGGAATGAGATCCAGTAATCGAGACTGACCTA 348  
OY 101 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 349 CAAGGCAATCCGAAATCGACTCAGATATTTAGCGGACCCCTCAAGTTTGCCTGTGAGAGCAT 408  
OY 121 ValGluGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 409 GTGGAGATATGAGAGATGAATCATTTGATTTCTCCAGAGAGCTGACAAATGTTAA 468  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
DB 469 GACAAACTTTCAGTATGACGAAACAGATTTGTGACCATGCGCTGCACATATCCCATAT 528  
OY 161 GluLeu 162  
DB 529 GAGCTA 534

RESULT 14  
AAK97837  
ID AAK97837 standard; cDNA: 832 BP.  
XX  
AC AAK97837:  
XX  
DT 23-SEP-1999 (first entry)  
XX  
XX Human secreted protein encoding cDNA #25.  
DE  
XX  
KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
KW diagnostic; gene therapy; chromosome mapping; secretion vector; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
PN W09925825-A2.  
XX  
PD 27-MAY-1999.  
XX  
PE 13-NOV-1998; 98WO-IB01862.  
XX  
PR 04-SEP-1998; 98US-0099273.  
PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.  
XX  
PA (GENY ) GENSET.  
XX  
PI Bouguieret L, Duclert A, Dumas Mline Edwards J;  
DR WPI: 1999-347472/29.  
DR P-PSDB; AAY36153.  
XX  
XX Extended cDNAs encoding secreted proteins  
PT  
XX  
PS Example 28; Page 194-195; 307pp; English.  
XX



CC PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.

XX Sequence 1210 BP; 261 A; 316 C; 379 G; 254 T; 0 other;

## Alignment Scores:

Pred. No.:	8,51e-91	Length:	1210
Score:	830.00	Matches:	158
Percent Similarity:	98.77%	Conservative:	2
Best Local Similarity:	97.53%	Mismatches:	2
Query Match:	97.99%	Indels:	0
DB:	22	Gaps:	0

US-10-082-502-21 (1-162) x AAS21330 (1-1210)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
DB 603 CGAGAGAGCCAGATCTCCCTGTGGAGCATGAGGGCTCTGTGGATGAACTAGAAATGG 662
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
DB 663 GAAATTGCCAGGTGACCCCAAGAACCATTCAGATGGGATCTTCCGGATCAATCCA 722
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 723 GATGGACGCCAGTCAGTGGTGAGTGCCCTTATGCCGCTCAGAGGCCCACTCAGAGAG 782
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
DB 783 CTCTGTGAGAGATATGTATGTACCGGATGAGAGATATGGGACAGATGTATCTTCACC 842
OY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 100
DB 843 CATCCCAAGAACTACGTACGTAGTGGCGCGGAATGAGAAATCCAGTGAATCGACCTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCATCCGAATCGACTAGATATTAGGGGACCTCAAGTTTCGTGTGAGACATT 962
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgLysLysAspAsnValLys 140
DB 963 GTGAGAGATACAGAGATGAATCATTTCTTTCCGAGAGGCTGCAATGTTTAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAAACTTTGACGTAAAGCAACAGATCTTTGTGACCATGCCCCGTGACATATCGCATGAT 1082
OY 161 GluLeu 162
DB 1083 GAGCTA 1088
RESULT 12
AAX97884
ID AAX97884 standard; cDNA: 714 BP.
AC AAX97884:
XX 23-SEP-1999 (first entry)
XX Human secreted protein encoding cDNA #72.
XX Secretd protein; human; cytosolic; thrombotic; osteopathic; forensic;
XX diagnostic; gene therapy; chromosome mapping; secretion vector; ss.
XX Homo sapiens.
XX MO9925825-A2.
XX 27-MAY-1999.
XX 13-NOV-1998; 98MO-IB01862.
XX PF 04-SEP-1998; 98US-0099273.
XX PR
```

PR 13-NOV-1997; 97US-0066677.  
PR 17-DEC-1997; 97US-0069957.  
PR 09-FEB-1998; 98US-0074121.  
PR 13-APR-1998; 98US-0081563.  
PR 10-AUG-1998; 98US-0096116.

XX (GEST ) GENSET.

XX Bouqueleret L, Duclert A, Dumas Mline Edwards J;

XX WPI: 1999-347472/29.

XX F-PSDB; AAT36200.

XX Extended cDNAs encoding secreted proteins

XX Claim 1: Page 272; 307pp; English.

CC AAX97813-X97906 represent extended cDNA's which encode novel human  
CC secreted proteins (see AAY36129-Y36222) and which have cytosolic,  
CC thrombotic and osteopathic activity. The extended cDNAs can be used to  
CC express secreted proteins or parts of them or to obtain antibodies  
CC capable of binding to the secreted proteins. They may also be used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC uses also include design of expression vectors and secretion vectors.

XX Sequence 714 BP; 203 A; 158 C; 202 G; 151 T; 0 other;

## Alignment Scores:

Pred. No.:	2,98e-90	Length:	714
Score:	823.00	Matches:	157
Percent Similarity:	98.15%	Conservative:	3
Best Local Similarity:	96.91%	Mismatches:	2
Query Match:	97.17%	Indels:	0
DB:	20	Gaps:	0

US-10-082-502-21 (1-162) x AAX97884 (1-714)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
DB 93 CGAGAGAGCCAGATCTCCCTGTGGAGCATGAGGGCTCTGTGGATGAACTAGAAATGG 152
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
DB 153 GAAATTGCCAGGTGACCCCAAGAACCATTCAGATGGGATCTTCCGGATCAATCCA 212
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
DB 213 GATGGACGCCAGTCAGTGGTGAGTGCCCTTATGCCGCTCAGAGGCCCACTCAGAGAG 272
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
DB 273 CTCTGTGAGAGATATGTATGTACCGGATGAGAGATATGGGACAGATGTATCTTCACC 332
OY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerSerGluLeuAspLeu 100
DB 333 CATCCCAAGAACTACGTACGTAGTGGCGCGGAATGAGAAATCCAGTGAATCGACCTA 392
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 393 CAAGGCATCCGAATCGACTAGATATTAGGGGACCTCAAGTTTCGTGTGAGACATT 452
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgLysLysAspAsnValLys 140
DB 453 GTGAGAGATACAGAGATGAATCATTTCTTTCCGAGAGGCTGCAATGTTTAA 512
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 513 GACAAACTTTGACGTAAAGCAACAGATCTTTGTGACCATGCCCCGTGACATATCGCATGAT 572
OY 161 GluLeu 162
DB 573 GAGCTA 578
```

CC polynucleotides, based on which tissues they are most highly expressed  
 CC in, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorders, immune deficiency diseases  
 CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC haematopoietic disorders, neural disorders, skeletal disorders,  
 CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The present sequence represents a  
 CC gene encoding a human secreted protein (see descriptor line for gene  
 CC number and clone identification).

XX Sequence 1180 BP; 258 A; 304 C; 363 G; 250 T; 5 other;

SO Alignment Scores:

Pred. No.: 8,24e-91 Length: 1180  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 0  
 Query Match: 97.99% Indels: 0  
 Gaps: 0

US-10-082-502-21 (1-162) x AAX22112 (1-1180)

QY 1 ArgArgSerGlnAAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 Db 573 CGAGAGACCCAGATCTCCACTGTGACAGCATGACAGGCGCTGTGATGAGTACATGAG 632  
 QY 21 GluIleAlaArgValAspProlyserThrIleGlnMetGlySerPheArgIleAsnPro 40  
 Db 633 GAATATGCCAGGTGAGCCCAAGAAACCATTCAGATGGATCTTCCGATCATCAATCA 692  
 QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 Db 693 GATGGACACCATGATGATGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 752  
 QY 61 LeuLeuGluGluValCysAspArgMetGlyGluTyrGlyGluGlnIleAspProSerThr 80  
 Db 753 CTGCTGAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 812  
 QY 81 HisArgLysAsnTyrValAlaGluValSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
 Db 813 CATGCCAAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 872  
 QY 101 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 Db 873 CAAGGCATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 932  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 Db 933 GTGGAGGAAATACGAGGAGGAGCACTCATGTAATCTTTCCGAGAGCGCAGCAATGTAAA 992  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 Db 993 GACAAACTTTCAGATGACGAGATCTTGTGACCATGCGCTGCACATATATGATGAT 1052  
 QY 161 GluLeu 162  
 Db 1053 GAGCTA 1058  
 RESULT 11  
 AAS21330  
 ID AAS21330 standard; cDNA; 1210 BP.  
 XX  
 AC AAS21330;  
 XX  
 DT 24-OCF-2001 (first entry)  
 DE Human cDNA sequence encoding for PRO4426 polypeptide.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;

KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIa; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO20010466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000WO-US32678.  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,  
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI: 2001-408281/43.  
 DR P-PSDB; AAU12258.  
 XX  
 FT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 FT other PRO polypeptides, link bioactive molecules to cells expressing  
 FT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 PT  
 XX  
 PS Claim 3; Fig 173; 813pp; English.  
 XX  
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding

PN	WO200190154-A2.
XX	29-NOV-2001.
PD	
XX	
PF	23-MAY-2001; 2001WO-US16695.
XX	
PR	24-MAY-2000; 2000US-207107P.
PR	13-JUN-2000; 2000US-211457P.
PR	21-JUN-2000; 2000US-213673P.
PR	03-AUG-2000; 2000US-223288P.
PR	01-MAR-2001; 2001US-272790P.
XX	
PA	(CORI-) CORIXA CORP.
PI	
PI	Xu J, Mitcham JL, Harlocker SL, Dillon DC, Secrist H, Lodes MJ;
PI	Algate PA, Fling SP, Mannion J, Benson DR, Carter D;
XX	
DR	WPI; 2002-097641/13.
XX	
PT	New isolated polynucleotide encoding polypeptide comprising portion of
PT	ovarian tumour protein, useful for detection, diagnosis and therapy of
PT	human ovarian cancer -
PS	Claim 1; Page 253-254; 285pp: English.
XX	
CC	The invention relates to an isolated polynucleotide encoding a
CC	polypeptide comprising a portion of an ovarian tumour protein. The
CC	sequences of the invention are useful for stimulating an immune response
CC	and for treating ovarian cancer in a patient. An antigen presenting cell
CC	that expresses the sequences is useful for treating ovarian cancer by
CC	incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells
CC	can then be proliferated and administered to the patient to inhibit the
CC	development of cancer. The DNA sequences are useful as probes or primers
CC	for nucleic acid hybridisation, to direct expression of a polypeptide in
CC	appropriate host cells. Detecting the presence of a cancer in a patient
CC	involves obtaining a biological sample from the patient, detecting the
CC	biological sample with an agent that binds to the protein, detecting the
CC	amount of protein that binds to the agent, comparing the amount of
CC	protein to a predetermined cut-off value and determining the presence of
CC	cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA
CC	molecules encoding ovarian tumour proteins of the invention.
XX	
XX	Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;
SO	
Alignment Scores:	
Pred. No.:	5.03e-91
Score:	830.00
Percent Similarity:	98.77%
Best Local Similarity:	97.53%
Query Match:	97.99%
DB:	24
US-10-082-502-21 (1-162) x ABK09772 (1-814)	
OY	1
DB	205
OY	21
DB	265
OY	41
DB	325
OY	61
DB	385
OY	81
DB	445

Oy	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle	120
Db	505	CAAGGATCCGATTCAGATCAGATATTTAGCGGACCCCTAAATTTGGCTGTGAGGCATT	564
Oy	121	ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys	140
Db	565	GTCGAGAAATACGAGCATGATGAACTCATTTGATTTCCGAGAGGCTGACAAATGTAA	624
Oy	141	AspLysLeuGlySerLysArgThrAspLeuGlyAspHisAlaLeuHisArgSerHisAsp	160
Oy	161	GACAACTCTTGGCAGTAAGCAGACAGATCTTTGTGACCAATGCCCTGCACATATGCATGAT	684
Db	685	GAGCTA 690	
DE	RESULT 10		
ID	AAK22112		
XX	AAK22112 standard; DNA; 1180 BP.		
AC	AAK22112;		
DT	18-MAY-1999 (first entry)		
XX	Human secreted protein gene 2 clone H2MB56.		
DE			
XX	Human; secreted protein; gene therapy; protein therapy; tissue; cancer;		
KW	leukaemia; neurodegenerative disorder; leukemias; autoimmune disease; AIDS;		
KW	developmental abnormality; foetal deficiency; Alzheimer's disease;		
KW	cognitive disorder; schizophrenia; immunological disorder; mood disorder;		
KW	immune deficiency disease; respiratory disorder; arthritis; skeletal;		
KW	haematopoietic disorder; neural; osteoporosis; metabolic disorders;		
KW	cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	W09901020-A2.		
XX			
PD	14-JAN-1999.		
XX			
PE	30-JUN-1998; 98MO-US13608.		
XX			
PR	12-SEP-1997; 97US-0058663.		
PR	01-JUL-1997; 97US-0051381.		
PR	01-JUL-1997; 97US-0051480.		
PR	12-SEP-1997; 97US-0058598.		
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;		
DR	WPI: 1999-105683/09.		
XX	P-PSDB; AAY01136, AAY01162, AAY01163.		
XX			
PT	New isolated human genes and the secreted polypeptides they encode -		
PT	useful for diagnosis and treatment of e.g. cancers, neurological		
PT	disorders, immune diseases, immune deficiency diseases or blood		
PT	disorders		
XX			
PS	Claim 4; Page 117, 179pp; English.		
XX			
CC	The invention relates to nucleic acid sequences (AAK22111 to AAK22134)		
CC	encoding human secreted proteins (AAY01135 to AAY01158). The secreted		
CC	protein gene sequences are deposited with the ATCC under deposit number		
CC	ATCC 209118. Host cells comprising recombinant vectors containing the		
CC	nucleic acid sequences are used for the recombinant production of the		
CC	secreted proteins. The polynucleotide and amino acid sequences are useful		
CC	for are useful for preventing, treating or ameliorating medical		
CC	conditions e.g. by protein or gene therapy. Pathological conditions can		
CC	be also diagnosed by determining the amount of the new polypeptides in a		
CC	sample or by determining the presence of mutations in the new		
CC	polynucleotides. Specific uses are described for each of the		

```
|||||
Db 265 GAAATGGCCAGAGTGGACCCCAAGAACCATTCAGATGGGATCTTTCGGATCATCA 324
Oy 41 AspGlySerGlnSerValValGluValProTyraIaArgSerGluAlaHisLeuThrglu 60
Db 325 GATGGCAGCAGCATGAGTGGTGGAGGTGCTTATGCCCGCTGAGAGCCACCTCACAGAG 384
Oy 61 LeuLeuGluGluValCysAspArgMetCysGluTyrglyGluGlnIleAspProSerThr 80
Db 385 CTGCTGGAGAGATATGTGACCGGATGAAGAGTATGGGGAACACATGATCTTCCACC 444
Oy 81 HisArgLysAsnItyrValArgValValSerItyrGlnGluSerSerGluLeuAspLeu 100
Db 445 CATCCCAAGAACTACGTACGTAGTACGTGGCCGGAATGGAGATCCAGTGAACCTGA 504
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 505 CAAGGCATCCGAAATGAGACTAGATATATGAGCGCACCCCAAGTTGGCGGTGAGAGCAT 564
Oy 121 ValGluGluTyrglyLysPheGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 565 GTGGAGGAAATACGAGAGTGAACCTCATTTCTTTCCCGAGAGCTGACAAATGTTAAA 624
Oy 141 AspLysIleCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
Db 625 GACAAACTTTGCAGTAAAGCAAGATCTTTGTGTACCATGCCCTGCACATGATGAT 684
Oy 161 GluLeu 162
Db 685 GAGCTA 690

RESULT 8
ABL41995
ID ABL41995 standard; DNA; 814 BP.
XX
XX ABL41995;
XX
Dl 11-JUN-2002 (first entry)
XX
De Nucleotide sequence of human polypeptide HP10390.
XX
KW Antibody; antigen; transmembrane domain protein; HP10390; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 145..693
FT /tag= a
FT /product= "HP10390"
XX
XX WO200208416-A1.
XX
XX 31-JAN-2002.
XX
XX 24-JUL-2001; 2001WO-JP06371.
XX
XX 24-JUL-2000; 2000JP-0222743.
XX
XX 24-AUG-2000; 2000JP-0254407.
XX
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
XX Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y;
XX
XX WPI; 2002-195877/25.
XX
XX P-PSDB; ABB09717.
XX
XX Antibody preparation by inoculation of an animal with a vector
XX expressing a fusion protein of an antigen on the C-terminal side of a
XX transmembrane domain for use as drugs, diagnostic reagents and
XX laboratory reagents
XX
XX Example; Page 33-35; 45pp; Japanese.
XX
```

```
CC The specification describes a method of antibody preparation. The
CC method comprises inoculating an animal with a vector expressing
CC a fusion protein having an antigen protein fused to the C-terminal
CC side (extracellular) of a transmembrane domain protein (the
CC N-terminal side of which is intracellular), and then isolating and
CC purifying the antibody from the animal. The antibodies can be used
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The
CC present sequence encodes a polypeptide, designated HP10390, which was
CC used in the course of the invention.
XX
SQ Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;
Alignment Scores:
Pred. No.: 5.03e-91 Length: 814
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 24 Gaps: 0
US-10-082-502-21 (1-162) x ABL41995 (1-814)
Oy 1 ArgArgSerGlnAspLeuHisCysGlyValAcysArgAlaLeuValAspGluLeuGlnTrp 20
Db 205 CGAGAGACCCAGAGATCTTCACCTGTGACACATGACAGGCTCTGGTGAATGACATGAATGG 264
Oy 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40
Db 265 GAAATGGCCAGAGTGGACCCCAAGAACCATTCAGATGGATGATCTTTCGGATCAATCA 324
Oy 41 AspGlySerGlnSerValValGluValProTyraIaArgSerGluAlaHisLeuThrglu 60
Db 325 GATGGCAGCAGCATGAGTGGTGGAGGTGCTTATGCCCGCTGAGAGCCACCTCACAGAG 384
Oy 61 LeuLeuGluGluValCysAspArgMetCysGluTyrglyGluGlnIleAspProSerThr 80
Db 385 CTGCTGGAGAGATATGTGACCGGATGAAGAGTATGGGGAACACATGATGATCTTCCACC 444
Oy 81 HisArgLysAsnItyrValArgValValSerItyrGlnGluSerSerGluLeuAspLeu 100
Db 445 CATCCCAAGAACTACGTACGTAGTACGTGGCCGGAATGGAGATCCAGTGAACCTGA 504
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 505 CAAGGCATCCGAAATGAGACTAGATATATGAGCGCACCCCAAGTTGGCGGTGAGAGCAT 564
Oy 121 ValGluGluTyrglyLysPheGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 565 GTGGAGGAAATACGAGAGTGAACCTCATTTCTTTCCCGAGAGCTGACAAATGTTAAA 624
Oy 141 AspLysIleCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
Db 625 GACAAACTTTGCAGTAAAGCAAGATCTTTGTGTACCATGCCCTGCACATGATGAT 684
Oy 161 GluLeu 162
Db 685 GAGCTA 690

RESULT 9
ABK09772
ID ABK09772 standard; cDNA; 814 BP.
XX
XX ABK09772;
XX
XX 14-MAR-2002 (first entry)
XX
XX Human ovarian tumour protein encoding cDNA #305.
XX
XX Human ovarian tumour protein; cancer; cytostatic; immunostimulant; ss;
XX gene therapy; CD4+ T cell; CD8+ T cell; PCR primer.
XX
XX Homo sapiens.
XX
```

CC which encodes a 20 kd protein with one putative transmembrane  
 CC domain in the N-terminus. The cDNA was isolated from a human stomach  
 CC cancer cell line cDNA library. The protein has no homology with any  
 CC known protein. The protein may be used to raise specific antibodies, as  
 CC assay reagents, as diagnostic tissue markers, for the isolation of  
 CC cognate receptors, ligands and binding proteins, and as biologically  
 CC active agents. Nucleotides encoding the protein may be used as primers  
 CC and probes or antisense molecules, and in gene therapy. Cells transformed  
 CC with these nucleotides may be used to screen for agonists and antagonists  
 CC which are potentially useful therapeutically.

SO Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:  
 Pred. No.: 5,03e-91 Length: 814  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 0  
 Query Match: 97.99% Indels: 0  
 DB: 21 Gaps: 0

US-10-082-502-21 (1-162) x AMZ38327 (1-814)

QY 1 ArgATGSeGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuTTP 20  
 DB 205 CGAGAGAGCCAGATCTCCACTGTGAGAGCATGCGAGGCTTGTGTGATGACATGATG 264  
 QY 21 GluTleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 DB 265 GAATATGCCAGGTGACCCCAAGAACATTCACATGGGATCTTCCGATCAATCCA 324  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 DB 325 GATGGAGCCAGCATGTGGAGGAGTCCCTTATGCGCCGCTCAGAGGCCACCTCACAG 384  
 QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 385 CTCTCTGAGAGCATATGTGACCGGATGAGAGATGGGAGACAGATGATCCTCCACC 444  
 QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 100  
 DB 445 CATCGCAAGAACTACGACTGTAGTGGCCGGAATGGAAATCCAGTCAATCGACCTA 504  
 QY 101 GlnGlyTleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 DB 505 CAAGGATCCGAAATCGACTGATATTAGGCGACCTCAAGTTGGCTGTGAGAGCAT 564  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 DB 565 GTGGAGAGATACAGAGATGAACTCATTTGTTCCCGAGAGGCTGACAAATGTTAAA 624  
 QY 141 AspLysLeuLysSerLysArgThrAspLeuCysAspHisAlaLeuHisAspSerHisAsp 160  
 DB 625 GACAAACTTTGCAGTAAAGCAACAGATCTTTGACCATTCCTGCACATATCGCATGAT 684  
 QY 161 GluLeu 162  
 DB 685 GAGCTA 690

RESULT 7  
 ID ABR52765  
 ID ABR52765 standard; cDNA; 814 BP.  
 AC ABR52765;  
 XX  
 XX  
 DT 15-AUG-2002 (first entry)  
 XX  
 XX cDNA encoding transmembrane protein 4, a cancer-linked protein.  
 DE Expressed sequence tag; EST; human; cancer; anti-neoplastic;  
 XX cytotoxic T lymphocyte; chemotherapy; cytostatic; chromosome 12;  
 KM transmembrane protein 4; gene; ss.  
 XX

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 145..693

FT /tag "a /product- "Transmembrane protein 4"

PN MO200231198-A2.

PD 18-APR-2002.

PE 11-OCT-2001; 2001MO-US31607.

PR 11-OCT-2000; 2000US-239294P.

PR 11-OCT-2000; 2000US-239297P.

PR 11-OCT-2000; 2000US-239605P.

PR 12-OCT-2000; 2000US-239802P.

PR 12-OCT-2000; 2000US-239805P.

PR 12-OCT-2000; 2000US-240622P.

PR 16-OCT-2000; 2000US-241682P.

PR 19-OCT-2000; 2000US-241723P.

PR 19-OCT-2000; 2000US-241723P.

PR 31-OCT-2000; 2000US-244932P.

(AVAL-) AVALON PHARM.

Young PE, Horrigan S, Weaver Z, Endress GA;

WPI; 2002-463271/49.

P-PSDB; AA097063.

Identifying modulators of a cancer-related gene to screen agents for  
 preventing or treating cancer comprises detecting a difference in the  
 expression of cancer-linked genes in the presence or absence of test  
 compounds

Claim 1; Page 43; 66pp; English.

The invention relates to modulators of a cancer-related genes. Also  
 described are: (1) processes for identifying an anti-neoplastic agents  
 comprising contacting a cell exhibiting neoplastic activity with a  
 compound first identified as a cancer related gene modulator, and  
 detecting a decrease in the neoplastic activity. (2) a process for  
 determining the cancerous state of a cell by determining an increase in  
 the level of expression of at least one gene, where an elevated  
 expression relative to a known non-cancerous cell indicates a cancerous  
 state or potentially cancerous state. The anti-neoplastic agent is  
 useful for treating cancer or for protecting an animal against cancer.  
 The immunogenic composition is also useful for treating cancer in an  
 animal, where the composition elicits the production of cytotoxic T  
 lymphocytes specific for the immunogenic composition. Preferably, the  
 animal is a human. The cancer-linked genes and polypeptides are also  
 useful as targets for cancer therapy or chemotherapy. The present  
 sequence represents a cancer-linked gene located on chromosome 12,  
 which encodes transmembrane protein 4.

SO Sequence 814 BP; 210 A; 194 C; 241 G; 169 T; 0 other;

Alignment Scores:  
 Pred. No.: 5,03e-91 Length: 814  
 Score: 830.00 Matches: 158  
 Percent Similarity: 98.77% Conservative: 2  
 Best Local Similarity: 97.53% Mismatches: 0  
 Query Match: 97.99% Indels: 0  
 DB: 24 Gaps: 0

US-10-082-502-21 (1-162) x ABR52765 (1-814)

QY 1 ArgATGSeGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGluLeuTTP 20  
 DB 205 CGAGAGAGCCAGATCTCCACTGTGAGAGCATGCGAGGCTTGTGTGATGACATGATG 264  
 QY 21 GluTleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40



Db 301 CATCGAAGACTACGTAGTGTGGCCGGAATGAGAAATCCAGTGAACCTGACCTA 360  
Oy 101 GINGLYILEARGYLEASPERASPILESERGLYTHLEULYSPHEALACYSGLUSERILE 120  
Db 361 CAAGGATCCGAAATGACATCAGATATTAGCGGACCCCTCAAGTTTGGTGTGAGACATT 420  
Oy 121 VALGLGLUTYRGILUASPGULUULEIIEGLUPHESERARGILUALAASPAENVALLYS 140  
Db 421 GTGCGAGCAATACGAGATCAACTGAAATCTTTCCCGAGAGGCTGCACATGTTAAA 480  
Oy 141 ASPLYSLEUCYSERLYSARGTHRASPLEUCYASPHISALALUHSARGSERHISASP 160  
Db 481 GACAACTTTGCGATGAAGCAACAGATCTTTGTGACCATGCTCCGACATATGCAATGAT 540  
Oy 161 GLULEU 162  
Db 541 GAGCTA 546  
RESULT 4  
AA06969  
ID AA06969 standard; cDNA; 806 BP.  
XX  
XX AAX06969;  
AC  
XX 10-MAY-1999 (first entry)  
DT  
XX Human secretory peptide-9 (Zs1g9) variant cDNA.  
DE  
XX Secretory peptide-9; Zs1g9; human; tumour marker; cancer; therapy;  
KM diagnosis; growth enhancer; variant; ss.  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH CDS 104..652  
FT /\*tag= a  
FT sig\_peptide 104..163  
FT /\*tag= b  
FT mat\_peptide 164..649  
FT /\*tag= c  
XX  
XX W09901554-A1.  
PN  
XX 14-JAN-1999.  
PD  
XX  
XX 02-JUL-1998; 98WO-US13859.  
PE  
XX  
XX 17-JUN-1998; 98US-0099005.  
PR 03-JUL-1997; 97US-0051704.  
PR 03-JUL-1997; 97US-0888088.  
PR 19-MAY-1998; 98US-0081338.  
PR 19-MAY-1998; 98US-0085983.  
PR 17-JUN-1998; 98US-0089899.  
XX  
XX (ZYMO) ZYMOGENETICS INC.  
PA  
XX Jaspers SR, Jelinek LJ, Sheppard PO, Whitmore TE;  
PI  
XX WPI: 1999-106055/09.  
DR P-PSDB: AAW88474.  
XX  
XX New mammalian secretory peptide-9 (Zs1g9) - used as a growth  
PT enhancer for placenta, liver and heart, and as an indicator of  
XX cancer  
XX  
XX Claim 2: Page 73-74; 85pp; English.  
XX  
XX This cDNA clone encodes human secretory peptide-9, or Zs1g9,  
CC variant (see AAW88474). Zs1g9 (see also AAW88469) is overexpressed in  
CC human brain, liver, lung, oesophageal, stomach, colon, rectal,  
CC thyroid and lymphoma tumors. Thus, Zs1g can be used as an  
CC indicator for cancer. Zs1g9 cDNA was discovered in a placenta  
CC clone from a full-term pregnancy cDNA library which contained an

CC expressed sequence tag (see AAX06971). The invention provides  
CC polynucleotides (see AAX06968-70) encoding Zs1g9 polypeptides (see  
CC AAW88469-77) including mature polypeptides, other processed forms,  
CC variants and mouse orthologues. The Zs1g9 gene, or probes derived  
CC from it, can be used to determine if Zs1g9 is present on chromosome  
CC 12, and if a mutation has occurred. Antibodies raised against  
CC Zs1g9 can be used as diagnostic agents to determine the presence of  
CC Zs1g9, and thus the presence of cancer. They can also be labelled  
CC with radioisotopes or fused with toxins and used to treat tumours  
CC which overexpress Zs1g9. Antisense nucleotides derived from Zs1g9  
CC cDNA can also be used to inhibit the growth of tumour cells. Zs1g9  
CC proteins can be used to enhance the growth or development of the  
CC placenta, heart or liver.  
XX  
XX Sequence 806 BP; 229 A; 183 C; 232 G; 160 T; 2 other;  
SO  
Alignment Scores:  
Pred. No.: 4.97e-91 Length: 806  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: 20 Gaps: 0  
US-10-082-502-21 (1-162) x AAX06969 (1-806)  
Oy 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20  
Db 164 CGAGAGACCCAGATCTCCCTGTGAGCATGCAAGGCTTGTGTGATCACTAGAAATGG 223  
Oy 21 GlnIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
Db 224 GAATATGCCAGGTGGACCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 283  
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuTrpGlu 60  
Db 284 GATGGAGCCAGTCAGTGTGGAGTGCCCTTATGCCCCCTCAGAGGCCCTCACAGAG 343  
Oy 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr 80  
Db 344 CTGCTGGAGAGATATGTGACCCGATGAGAGATATGGGAGAACAGATGATCCTCCACC 403  
Oy 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100  
Db 404 CATCGCAAGAACTACGATGCTAGTGGCCGCGAATGAGAAATCCAGTGAACCTGACCTA 463  
Oy 101 GINGLYILEARGYLEASPERASPILESERGLYTHLEULYSPHEALACYSGLUSERILE 120  
Db 464 CAAGGATCCGAAATGACATCAGATATTAGCGGACCCCTCAAGTTTGGTGTGAGACATT 523  
Oy 121 VALGLGLUTYRGILUASPGULUULEIIEGLUPHESERARGILUALAASPAENVALLYS 140  
Db 524 GTGCGAGCAATACGAGATGACTGATTTGATTTTCCCGAAGGCTGCACATGTTAAA 583  
Oy 141 ASPLYSLEUCYSERLYSARGTHRASPLEUCYASPHISALALUHSARGSERHISASP 160  
Db 584 GACAACTTTGCGATGAAGCAACAGATCTTTGTGACCATGCTCCGACATATGCAATGAT 643  
Oy 161 GLULEU 162  
Db 644 GAGCTA 649  
RESULT 5  
AA08293  
ID AA08293 standard; DNA; 806 BP.  
XX  
XX AA08293;  
AC  
XX 07-FEB-2000 (first entry)  
DT  
XX Human Zs1g9 gene encoding secretory protein variant-4.  
DE  
XX Secretory protein-9; Human Zs1g9; chromosome 12q15 region; variant;  
KW



XX		The present DNA sequence is an ortholog encoding the secretory protein-9.
CC	Zs1g9	derived from mouse. It is overexpressed in tumours. Antagonists,
CC	antibodies and antisense nucleotides to Zs1g9 are useful for detecting	
CC	and treating tumours. The antagonist may be an antibody or receptor to	
CC	Zs1g9 and it may be radio-labelled or fused to a polypeptide toxin. It	
CC	can be used for down regulating the overexpression of Zs1g9. The gene	
CC	sequence can be used as nucleic acid probes to detect RNA encoding Zs1g9	
CC	The Zs1g9 sequence facilitates improved diagnostic and therapeutic	
CC	techniques for detecting and treating cancers, especially of the brain,	
CC	liver, stomach, lymphoma etc., at an early stage.	
XX		
SQ	Sequence	1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;
	Alignment Scores:	
	Pred. No.:	6,29e-93 Length: 1069
	Score:	847.00 Matches: 162
	Percent Similarity:	100.00% Conservative: 0
	Best local Similarity:	100.00% Mismatches: 0
	Query Match:	100.00% Indels: 0
	DB:	21 Gaps: 0
US-10-082-502-21	(1-162) x AA208294	(1-1069)
OY	1	ArgArSerGlnAspLeuHisCysGlyAlaCyArgAlaLeuValaspJluLeuGluTrp 20
Db	418	CANAGGAGCCCAAGTCTACACTGTGGAGCTTGAGGGCTCTGGTGATTAATTAAAGTTGG 477
OY	21	GluIleuAaArgValaAspProLysLythrIleGlmMetGlySerPheArgIleaAspro 40
Db	478	GAATTGCCCGCGGAGACCACCAAGAAGACCATTCACAATGGANTCCITCCGAATCATCCA 537
OY	41	AspelysSerGlnSerValValGlutValProTyrrAlaArgSerGluAlaHisLeuThrglu 60
Db	538	GATGGCAGCCAGTCAGTGTGGAGGATCACTTATGCCCGCTCAGAGGCCACCTCACAGAG 597
OY	61	LeuengGluGluValLysasParGmetLysGluTyrrGlyGlnGlnIleAsProSerThr 80
Db	598	TTCGTTGAGGAGGTGTGTGACCAATGAAGAGTAGAGGGAACAGATTGACCCCTTATACC 657
OY	81	HIsArGlySaAntyryValaArgValaIserArgAsnGlyGluSerSerGluLeuAspLeu 100
Db	658	CACCGCAAGAACATACGTAACGCCTGTGAGCGGGAATGGAGNATCCAGTAACTAGACTTA 717
OY	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerile 120
Db	718	CAGGCGATCCGAATTGACTCAGATATCAGCGGCAACCCCTCAAGTTGGCGTGAAGACATT 777
OY	121	ValGlnIuLuryrGluNaSpGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
Db	778	GTGGAAACAATRACGAGGTAGCTTATCGAATCTTCTCCAGAGAGGCTGACAAACGTTAAA 837
OY	141	AspLysLeucYseSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
Db	838	GACAAACTTTTGCAGTAAAGCGGACAGATCTATGTGACCAATGCCCTGCACAGATCTACGAT 897
OY	161	GluLeu 162
Db	898	GAGCTA 903
RESULT 3		
AAZ38326		
ID	AAZ38326	standard; cDNA; 546 BP.
AC	AAZ38326;	
XX		
DT	09-FEB-2000	(first entry)
XX		
DE	Human transmembrane protein cDNA clone HP10390 coding sequence.	
KW	HP10390; transmembrane domain; stomach cancer cell; antibody;	
KM	assay reagent; diagnostic marker; primer; probe; antisense; gene therapy;	
KW	agonist; antagonist; ligand; therapeutic; ds.	

XX	Homo sapiens.
OS	
XX	Key Location/Qualifiers
FT	CDS 1..540
FF	/lag- <sup>a</sup>
FM	/product= "human transmembrane protein HP10390"
FH	/note= "No stop codon given in the specification"
XX	
PX	M09J95586Z-A2.
PN	
PD	04-NOV-1999.
PF	
PE	27-APR-1999; 99MO-JR02226.
PR	28-APR-1998; 98JP-0119395.
PA	(SAGA ) SAGAMI CHEM RES CENT. (PROT-) PROTEGENE INC.
PI	Kato S, Kimura T;
DR	WPI: 2000-023358/02.
XX	P-PADB; AAY52391.
PT	Human proteins with transmembrane domains, involved in control of cell proliferation and differentiation, useful for treating e.g. cancer or inflammation -
PS	Claim 3; Page 88; 11App; English.
CC	This sequence represents the coding sequence of human cDNA clone HP10390 which encodes a 20 kD protein with one putative transmembrane domain in the N-terminalus. The cDNA was isolated from a human stomach cancer cell line cDNA library. The protein has no homology with any known protein. The protein may be used to raise specific antibodies, as assay reagents, as diagnostic tissue markers, for the isolation of cognate receptors, ligands and binding proteins, and, as biologically active agents. Nucleotides encoding the protein may be used as primers CC and probes or antisense molecules, and in gene therapy. Cells transformed with these nucleotides may be used to screen for agonists and antagonists which are potentially useful therapeutically.
SQ	Sequence 546 BP; 139 A; 126 C; 165 G; 116 T; 0 other;
Alignment Scores:	
Score:	2,96e-91 Length: 546
Percent Similarity:	830.00 Matches: 158
Best Local Similarity:	98.778 Conservative: 2
Query Match:	97.538 Mismatches: 0
DB:	Indels: 2
	Gaps: 0
D5-10-082-502-21 (1-162) x AA238126 (1-546)	
OY     1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGlutrp     20	
Dd       CGAGAGGCCACAGATCTCCACTGTGGAGCATGGAGGCCTGTGTGAAGAACTGAATTGG   120-	
OY     21 GluIleAalrGyValasPrOdylsyvTrllegImetglySerPheargtLleasnPro    40	
Dd       :::~::~~::~~::~~::~~::~~::~~::~~::~~::~~::~~::~~::~~::~~::~~::~~::~~::	
Dd     121 GAATTTCCCAAGGGCCACCACAAGAACACATTCAGATGGATCTTTCCGGATCAATCCA   180	
OY     41 AspelserGINserValvalIGluValproCyralArgrSerGIuaINhleuthrlu     60	
Dd       181 GATGGCACCCAGTCAGTGGTGGAGGCTGTATGCCCTCCAGAGCCCACCTCAACAG     240	
OY     61 LeuleudIGLuValaCysaspArmetylgUUTryGLygluInIlaspProsetThr     80	
Dd       241 CTGCTGAGGAGAtATATCTGACCGGATGAAGAGATATGGGAAACAATTAATCTTTCAACC   300	
OY     81 HisArgLyasanTyrrValArgValValSerArgAsnGlygIUsserSerGIuleuAspleu   100	

FT	sig_peptide	358..417
FT	/*tag= b	
FT	mat_peptide	418..903
FT	/*tag= c	
XX		
PN	MO9901554-A1.	
XX		
PD	14-JAN-1999.	
XX		
PE	02-JUL-1998;	98NO-US13859.
XX		
PR	17-JUN-1998;	98US-0099005.
XX		
PR	03-JUL-1997;	97US-0051704.
PR	03-JUL-1997;	97US-0888088.
PR	19-MAY-1998;	98US-0081338.
PR	19-MAY-1998;	98US-0085983.
PR	17-JUN-1998;	98US-0089899.
XX		
PA	(ZYMO ) ZYMOGENETICS INC.	
XX		
PI	Jaspers SR, Jellinek LJ, Sheppard PO, Whitmore TE;	
XX		
DR	WPI: 1999-106055/09.	
DR	P-PSDB: AAM88476.	
XX		
PT	New mammalian secretory peptide-9 (Zs19) - used as a growth	
PT	enhancer for placenta, liver and heart, and as an indicator of	
PT	cancer	
XX		
PS	Claim 2; Page 75-77; 85pp: English.	
XX		
CC	This cDNA clone encodes novel mouse secretory peptide-9, or Zs19	
CC	(see AAM88476), an orthologue of novel human Zs19 (see AAM88469).	
CC	human Zs19 is overexpressed in a number of tumours including	
CC	brain, liver, lung, oesophageal, stomach, colon, rectal, thyroid	
CC	and lymphoma tumors, and thus can be used as an indicator for	
CC	cancer. The invention provides polynucleotides (see AAM06568-70)	
CC	encoding Zs19 polypeptides (see AAM88469-77) including mature	
CC	polypeptides, other processed forms, variants and the mouse	
CC	orthologues. The Zs19 gene, or probes derived from it, can be	
CC	used to determine if Zs19 is present on chromosome 10, and if a	
CC	mutation has occurred. Antibodies raised against Zs19 can be	
CC	used as diagnostic agents to determine the presence of Zs19, and	
CC	thus the presence of cancer. They can also be labelled with	
CC	radioisotopes or fused with toxins and used to treat tumours	
CC	which overexpress Zs19. Antisense nucleotides derived from Zs19	
CC	cDNA can also be used to inhibit the growth of tumour cells. Zs19	
CC	proteins can be used to enhance the growth or development of the	
CC	placenta, heart or liver.	
XX		
S0	Sequence 1069 BP; 276 A; 259 C; 325 G; 209 T; 0 other;	
Alignment Scores:		
Pred. No.:	6,29e-93	Length: 1069
Score:	847.00	Matches: 162
Percent Similarity:	100.00%	Conservative: 0
Best local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	20	Gaps: 0
US-10-082-502-21 (1-162) x AAX06970 (1-1069)		
OY	1	ARGATGSEGTGAADPLeNHISySGSLyALAcysAATgGAlaLeuValAAsPGluLeuGluTrp 20
DB	418	CGAAGACCCAAAGATCTACATGTGGACCTTCAGGGCTCTGGTGGATGAATTAGAACTGG 477
OY	21	GLuLeaLaarGValAAsPProLyLaLyThrILleGInMeTgISerPheArgILeasPro 40
DB	478	GAATATGGCCCGCGTGGACCCCAAGAACCATTCAGATGGATCCCTCCGAATCAATCCA 537
OY	41	AspGLysErGInSerValValGluValProGTrgLaLaarGserGILuAlaHisLeuThGlu 60
DB	538	GATGGCACCCAGTGTGTGTGAGGTACCTTATCCGCGCTAGAGGCCACCTACAGAG 597

QY	61	LeuEngluGlutValCysAspArgMetLysGluTyrGluGluGlnIleAspProSerThr	80
Db	598	TTCGTTGAGAGAGTGTCTGACCGAATGAAGAGATACGGGAGAACAGATTCACCTTCTACC	657
QY	81	HisArgLysAsnTyrValArgValValSerArgAsnGlyGlnSerGluLeuAspLeu	100
Db	658	CACCGCAAGAACTACACTACGCGTGTGAGCCGAGATGGAGAACTGAACTAGACTTA	717
QY	101	GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheIleAcylGluSerIle	120
Db	718	CAGGCGCTTCGGAATTTACATCAATATACAGCGGCACCTCCAAATTGGCTGTGAGACATTT	777
QY	121	ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys	140
Db	778	GTGGAAACAAATACAGAGATGAGCTTATCCGAATTTCTTCCACAGAGCGCTGACAACTTAA	837
QY	141	AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp	160
Db	838	GACMAACTTTGGCAGTAAAGCGGACAGATCTATGTGACATGCCCTGCACAGATCTCACAGAT	897
QY	161	GluLeu 162	
Db	898	GAGCTA 903	
RESULT 2			
QY	AAZ08294	standard; DNA; 1069 BP.	
XX	AAZ08294;		
XX	07-FEB-2000	(first entry)	
XX	Mouse ortholog gene encoding Zs1g9 secretory protein.		
XX	Secretory protein-9; mouse Zs1g9; ortholog; overexpression; antagonist;		
KW	antibody; antisense nucleotide; tumour; treatment; receptor;		
KW	radio-labeled; polypeptide toxin; down-regulation; diagnostic;		
KW	therapeutic; probe; cancer; brain; liver; detection; stomach;		
KW	lymphoma; ds.		
OS	Mus musculus.		
XX	Key		
XX	Location/Qualifiers		
FT	CDS	358..906	
FT	/*tag= a		
FT	/product= "Mouse Zs1g9 secretory protein"		
FT	/note= "Overexpressed in tumours"		
FT	s1g_peptide	358..417	
FT	/*tag= b		
FT	mat_peptide	418..903	
FT	/*tag= c		
FT	/label= Mature_Zs1g9_protein_variant-4		
XX	W03960405-A1.		
XX	25-NOV-1999.		
XX	19-MAY-1999;	99MO-US11107.	
XX	19-MAY-1998;	98US-0081183.	
XX	(ZYMO ) ZYMOGENETICS INC.		
XX	Moore EE, Taft DW;		
XX	WPI; 2000-039447/03.		
XX	P-PSDB; AAY15136.		
XX	Detecting tumors using antibodies, antagonists and antisense		
XX	nucleotides to secretory protein-9 (Zs1g9)		
XX	Disclosure; Page 37-38; 45pp; English.		

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:08:03 ; Search time 127.151 Seconds  
(without alignments)  
2869.211 Million cell updates/sec

Title: US-10-082-502-21

Perfect score: 847  
Sequence: 1 RRSODLHCACALVDELEW.....LCSKRTDLCDAHLRSHDEL 162

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=framet+2pn.model -DEV=xip  
-Q=/cg21.1/USPTO.spool/US10082502/rnat.24032003.135058.6538/app.query.fasta.1.1308  
-DB=N\_Geneseq.101002 -GEMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIG=200 -THR=SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTEXT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USR=US10082502.ccgcn.1.1.741.grnat.24032003.135058.6538 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-MAIN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq.101002.\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	847	100.0	1069	20 AAX06970	Mouse secretory pe
2	847	100.0	1069	21 AAX08294	Mouse orthology gen
3	830	98.0	546	21 AAX238326	Human transmembran
4	830	98.0	806	20 AAX06969	Human transmembran
5	830	98.0	806	21 AAX08293	Human Zs19 gene e
6	830	98.0	814	21 AAX238327	Human transmembran
7	830	98.0	814	24 AAK52765	CDNA encoding trans
8	830	98.0	814	24 ABL14995	Nucleotide sequenc
9	830	98.0	814	24 ABR09772	Human ovarian tumo
10	830	98.0	1180	20 AAX22112	Human secreted pro
11	830	98.0	1210	22 AAX21330	Human secreted pro
12	823	97.2	714	20 AAX97884	Human secreted pro
13	815	96.2	657	24 ABR36007	CDNA sequence #398
14	781	92.2	832	20 AAX97837	Human secreted pro
15	763.5	90.1	592	21 AAX43525	Mouse secreted exp
16	380	44.9	484	22 AAL25628	Human breast cance
17	380	44.9	558	22 AAL18023	Human breast cance
18	380	44.9	595	22 AAL07882	Human breast cance
19	380	44.9	1160	22 AAL26689	Human breast cance
20	253	29.9	649	20 AAX97883	Human secreted pro
21	253	29.9	649	20 AAX06968	Human secreted pro
22	253	29.9	649	20 AAX06968	Human secreted pro
23	253	29.9	718	20 AAX208284	Human secreted pro
24	250	29.5	792	20 AAX97836	CDNA encoding a hu
25	249	29.4	415	20 AAX06971	Human secreted pro
26	249	29.4	415	21 AAX08285	Zs19 expressed sequ
27	239	28.2	1085	20 AAX22130	WO9010120 seq ID 3
28	204	24.1	816	23 ABL21471	Drosophila melanog
29	201	23.7	2532	22 ABA07341	Human pancreatic c
30	201	23.7	2532	22 AAS27268	Human genomic DNA
31	201	23.7	2724	22 ABA07340	Human pancreatic c
32	201	23.7	2724	22 AAS32767	Human genomic DNA
33	155	18.3	786	21 AAC37767	Ara1idopsis thalia
34	154	18.2	564	24 ABR09691	Human ovarian tumo
35	141.5	16.7	750	23 ABL16485	Drosophila melanog
36	141.5	16.7	2750	23 ABL16484	Drosophila melanog
37	141.5	16.7	3274	23 ABL16482	Drosophila melanog
38	140	16.5	3110	23 ABL21470	Drosophila melanog
39	132	15.6	744	21 AAX49678	Human myocardiolum s
40	132	15.6	746	22 AAF94104	Primer specific fo
41	132	15.6	747	22 AAC90704	Human secretory pr
42	132	15.6	1419	22 AAX93934	Human CDNA encodin
43	132	15.6	1457	20 AAX78923	Human zcalc-1 DNA
44	132	15.6	1462	21 AAL16621	Human secreted pro
45	132	15.6	1470	21 AAX96349	CDNA encoding a no

## ALIGNMENTS

RESULT 1  
AAX06970 standard; CDNA; 1069 BP.

AC AAX06970;

XX 10-MAY-1999 (first entry)

XX Mouse secretory peptide-9 (Zs19) orthologue CDNA.

XX Secretory peptide-9; Zs19; orthologue; mouse; tumour marker;

XX cancer; therapy; diagnosis; growth enhancer; ss.

XX Mus sp.

XX Key Location/Qualifiers

FT CDS 358..906

FT /\*tag- a



CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/139,491  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/455,525  
FILING DATE: 31-MAY-1995  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 601567and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 148..2910  
US-09-139-491-38

Alignment Scores:  
Pred. No.: 12.3 Length: 4131  
Score: 76.00 Matches: 37  
Percent Similarity: 37.06% Conservative: 36  
Best Local Similarity: 18.78% Mismatches: 68  
Query Match: 8.97% Indels: 56  
DB: 3 Gaps: 7

US-10-082-502-21 (1-162) x US-09-139-491-38 (1-4131)

OY 6 LeuHlscYsglyAlaCyArGAlaLeuValAspGluleuGluTrpGluIleAlaArgVal 25  
:::|||||  
DB 574 GTGCATCTGTGT-----CAGCTGAGTGCATGAGAGTGGAGCTGCACACTGTG 624  
OY 26 AsprPolys-----LysThrIleGlnMetGlySerPheArg 37  
:::|||||  
DB 625 GAGAAAGCATACCTGTGGCCCTGAAAGGCTGCAAGCCCTGCACAGCGGAGTCCAGC 684  
OY 38 IleAnpProAspGlySerGlnSerValValGlu----- 48  
:::|||||  
DB 685 GTGGCCCGGAGAGCGACCCAGAAATCTCCGGAGAGAGCGGCGGAGACAGCAAGGGTGGG 744  
OY 49 ValProTyrAlaArgSerGluAlaHisLeuThrGluLeuGluGluValCysAspArg 68  
||| ||| ::::|  
DB 745 GTGCGATACACAACCAAGACGAAAGATCTGCGACGCTTTGGGGGAGCTCTACGAC--- 801  
OY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValArgVal 88  
:::|||||  
DB 802 -----CTGATGATCATCTCTCCGACGCTCAAGCTCCTCAATAT 840  
OY 89 ValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArgIleAspSerAsp 108  
:::|||||  
DB 841 CTGCAACAGAGACCCAGGCAATCCGCTGCTGCTGCTGTATCCGAGGACAAT--- 897

OY 109 IleSerGlyThrLeuLysPheAlaCysGlnSerIleValGluGluTyrGluAspGluLeu 128  
|||:::|  
DB 898 -----CTTCAGCTCTCTCTGCAAGGTCATTTGAGATTAAGTACTGGAGAGAG 945  
OY 129 IleGluPhe-----PheSerArgGluAlaAspAsnValLysAspLysLeu----- 143  
||| ||| |||  
DB 946 ATCACTTTCGCTTGACCAAGAGCCTGGCCCAAGTGTGGAGACAAAGATCTATATC 1005  
OY 144 -----CysSer 145  
|||  
DB 1006 CAGCTGAAAGATCTACCTCCGAGATATGCAACAGCTGCAAAAGCATGTGGCTGTGAG 1065  
OY 146 LysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAspGluLeu 162  
:::|||||  
DB 1066 GTGAGGCCATGCTGTGTGCTCATCAAGCCGCGCACTGACCAAGGTC 1116

Search completed: March 31, 2003, 04:26:21  
Job time : 45.3198 secs







Db 1066 GTCCAGCCATGCTGTGTCTCCCTGTCATCCAGCGGGCCACTGACGAGTTC 1116  
RESULT 11  
US-08-297-510-38  
Sequence 38, Application US/08297510  
Patent No. 5602019  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/297,510  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5602019and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 346-5750  
TELEFAX: (312) 984-9740  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4131 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 148..2910  
US-08-297-510-38  
Alignment Scores:  
Pred. No.: 12.3 Length: 4131  
Score: 76.00 Matches: 37  
Percent Similarity: 37.06% Conservative: 36  
Best Local Similarity: 18.78% Mismatches: 68  
Query Match: 8.97% Indels: 56  
Gaps: 7  
US-10-082-502-21 (1-162) x US-08-297-510-38 (1-4131)  
QY 6 LeuNtSCySbLlYlAcSvAtGdAlaLeuValAspRlLeuNlUtrRlGluLleAlaVal 25  
Db 574 GTGACACTGTGT-----CAGCTGAGTGCACATGAGAGTGGAGCCTGCAAGCTGTG 624  
QY 26 AspRlGluS-----LysThrLleGlnMetGlySerPheArg 37  
Db 625 GAGAAAGATACCTGTCGTGGCCCTGAAAGAGCTGCAGGCTTGCAGCAGCGAGAGTCCAGC 684

QY 38 IleAsnProAspLysSerGlnSerValValGlu----- 48  
Db 685 GTGGCCCCCGGAAGCAGCACCAAGATCTCCGAGAGAGCAGCGGAGACCAAGAGGTGG 744  
QY 49 ValProTyrAlaArgSerGlnAlaHisLeuThrGluLeuGluGluValCysArg 68  
Db 745 GTCCGATCACCAACCAACCAAGATCTGACCTTGGCGGGAGCTTCACGAC 801  
QY 69 MetLysGluTyrGlyGluGlnIleAspProSerThrHisArgLysAsnTyrValAlaVal 88  
Db 802 -----CTGAGTCATCTTCCCTCAGCTCAAAGTCTCCCATAT 840  
QY 89 ValSerArgAsnGlyLysSerSerGluAspLeuGlnGlyIleArgLysSerAsp 108  
Db 841 CTCACACAGAGACCAAGGATCCCGCTGCTGCTGCTGCTGATCCAGAGCAT 897  
QY 109 IleSerGlyThrLeuLysPheAlaCysGluSerIleValGluGluTyrGluAspGluLeu 128  
Db 898 -----CTTCAGCTCTCTCCCAAGTCATTCAGATTAAGTACTGAGAGAG 945  
QY 129 IleGluPhe-----PheSerArgGluAlaAspAsnValLysAspLysLeu 143  
Db 946 ATCAGCTTCCGTTGACACACAGAGCGCTGGGCCAAAGTGTGGAACAGACATCTATC 1005  
QY 144 -----CysSer 145  
Db 1006 CAGCTGAAGATCTACCTCCGAGATATGCAACAGCTGCAAAGCATGTGGCTGTAG 1065  
QY 146 LysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAspGluLeu 162  
Db 1066 GTCCAGCCATGCTGTGTCTCCCTGTCATCCAGCGGGCCACTGACGAGTTC 1116  
RESULT 12  
US-08-479-532-38  
Sequence 38, Application US/08479532  
Patent No. 5776752  
GENERAL INFORMATION:  
APPLICANT: Beavo, Joseph A.  
APPLICANT: Bentley, Kelley  
APPLICANT: Charbonneau, Harry  
APPLICANT: Sonnenburg, William K.  
TITLE OF INVENTION: DNA Encoding Mammalian  
TITLE OF INVENTION: Phosphodiesterases  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
STREET: Two First National Plaza, 20 South Clark  
STREET: Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,532  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/297,494  
FILING DATE:  
APPLICATION NUMBER: US 07/688,356  
FILING DATE: 04-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5776752and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 27866/30822  
TELECOMMUNICATION INFORMATION:



```
Oy      145 SerLysargThrAspLeu 150
          ::::|:::  ::::|
Db      956 ACTAAAAAGCTGAATTG 973
```

## RESULT 8

Sequence 208, Application US/08961527  
Patent No. 6420135  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 391  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 208:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3752 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-208

Alignment Scores:	
Pred. No.:	9, 23
Score:	76.50
Percent Similarity:	39.77%
Best Local Similarity:	21.22%
Query Match:	9.03%
DB:	4
Length:	3752
Matches:	31
Conservative:	27
Mismatches:	65
Indels:	23
Gaps:	4

US-10-082-502-21 (1-162) x US-08-961-527-208 (1-3752)

```

Oy      15 Valaerpleuleugltupgluilaala-----ArgValaerProLysThr 30
      :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      701 ATTCTCTACTTTGGACAAGAAAGTTGCTACTGCTCAACATCAAGTAAAGTAATTTGGAAAAA 760
      :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy      31 IleglmetcglyserpheatrgileasnproaspelyserginserValaIgluValpro 50
      :: :: ||||| :: ||||| ||||| ||||| ||||| ||||| |||||
Db      761 CTCTCTGCTGCTGGC-----GATCCTGATGATGGACAGAAAGTTATAGAAAGCTAAA 811
      :: :: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy      51 TyrAlaArgSerGluAlaIAsLeuThrGluLeuGluGluValaCysAspArgMetLys 70
      :: :: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      812 TTAAAAAAGAGCAAGCACTGACCTAAAGCTTAAACACAGCTGAGTTACGAAAAAACAACA 871
      :: :: ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy      71 GluTyrGluGluGluIleAspProSerThrIleArgLysasnTyrValaIrgValaValser 90
      :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      872 GAACCTTGAAAAAATCTTGTGACAGCCCTTGATCCTCAAGTAAGACTCAGATGAATGAT 931
      :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY	91	Argmsnglueberseglueaspleingllylearglueaspeaserpleer	110
Db	932	AAAGAACCAAGACAGCTGAGTGGATATAAAAAGCGATGAACCTCAAAATTAAGTCTCT	991
QY	111	GLyThrleuylsPhealacygluserllevalglutyrGlualspcluleuileglu	130
Db	992	GATTTA-----GAAAAGAATAATGTAACCTTGAA	1021
QY	131	Pheheserargglunlaasp-----Asnvallyasplylsleucys	144
Db	1022	ATATTACTTGGAAGGCGTGAATCCTGAAGATGATACTGCTGCTTCATCAAAATTAATTAAGCT	1081
QY	145	SerlysarGrThraspleu	150
Db	1082	GCTAAAAAGACTGAGTTA	1099

## RESULT 5

```

1 Sequence 38, Application US/07872644
2 Patent No. 5389527
3
4 GENERAL INFORMATION:
5
6 APPLICANT: Beavo, Joseph A.
7
8 APPLICANT: Bentley, Kelsey
9
10 APPLICANT: Chabonneau, Harry
11
12 APPLICANT: Sonnenburg, William K.
13
14 TITLE OF INVENTION: DNA Encoding Mammalian
15
16 Phosphodiesterases
17
18 NUMBER OF SEQUENCES: 58
19
20 CORRESPONDENCE ADDRESS:
21
22 ADDRESSEE: Marshball, O'Toole, Gerstein, Murray &
23
24 ADDRESSEE: Bicknell
25
26 STREET: Two First National Plaza, 20 South Clark
27
28 STREET: Street
29
30 CITY: Chicago
31
32 STATE: Illinois
33
34 COUNTRY: USA
35
36

```

Alignment Scores:	
Pred. No.:	12.3
Score:	76.00
Percent Similarity:	37.068
Length:	413
Matches:	37
Conservative:	36

```
TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6519 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL:
: ANTI-SENSE:
: FRAGMENT TYPE:
: ORIGINAL SOURCE: spleen cell of homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 24..5619
: US-08-971-988-1

Alignment Scores:
Pred. No.: 13.1 Length: 6519
Score: 78.00 Matches: 44
Percent Similarity: 40.11% Conservative: 29
Best Local Similarity: 24.18% Mismatches: 66
Query Match: 9.21% Indels: 44
DB: 1 Gaps: 8

US-10-082-502-21 (1-162) x US-08-971-988-1 (1-6519)
OY 6 LeuHtscysgIyAlaCySArgAlaLeuValAspLueuGIuTPrGluIleAlaArgVal 25
Db 4660 TTACAGACCGGTACTGTCAGAGACACCCCTGAGGCCCATGAAA-----AGATCG 4707
OY 26 AspProLysIyThrIleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSer 45
Db 4708 AGAAGCTCAAGAGCATGATCTGCGCAGATTCCTTTCTGCGCAGAGGATCAGAAATCC 4767
OY 46 ValValGIuValProTyRAlaArgSerGIuAlaHisLeuThrGIuLeuLeuGIuGIuVal 65
Db 4768 ATG-----GAGACAAA-GTCACGGAGGACCTGAGGCCGTTTC 4802
OY 66 CysAspArgMet-----LysGIuTyRGIy 73
Db 4803 CACGAGAGGATGAGGAGCCCTGTTTCAACACAGCTGAAGGAAAAAGTGGACAAAGATACGCC 4862
OY 74 GIuGlnIleAspProSer-----ThrHisArgLysAsnTyRVal 86
Db 4863 GTCCGAATCATGCCCTCAAGTCTGTGATGATGAAGAGGACGCCGCCCGGTCATGCGTTC 4922
OY 87 ArgValAlaSerArgAsnGlyGlySerSerGIuLeuAspLeuGlnGlyIle---ArgIle 105
Db 4923 CGGTCTTCACAGATGCCCTTCTCATCCGCCCTGTGTGTGGCCCTGTCTTCCCTTC 4982
OY 106 AspSerAspIleSer-----GlyThrLeuLysPheAlaCySGLuSerIleVal 121
Db 4983 TCATCGGAGACAGCACCCCTCCAGACAGCAGCCGCGGTTTCCCTCGACCCCTCTCTCG 5042
OY 122 GIuGIuTyRGIuAspGIuLeuIleGlnPhePheSerArgGIuAlaAspAsnVal---Lys 140
Db 5043 CCAAG-----AAATGCATCTCAGAGTCCAGTCCACACAACTGACACAG 5084
OY 141 AspLysLeuCySerLysArgThrAspLeuCySAspHisAlaLeuHisArgSerHisAsp 160
Db 5085 GATACCTGGAGAAAGAGAAAGACAAAGAAAGAAAAAGAACACAGCAACATCAAA 5144
OY 161 GIuLeu 162
Db 5145 GAGATA 5150

RESULT 7
US-08-961-083-159
: Sequence 159, Application US/08961083
: Patent No. 6159469
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
```

```
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
: NUMBER OF SEQUENCES: 432
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8512
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 159:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1924 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-961-083-159

Alignment Scores:
Pred. No.: 2.64 Length: 1924
Score: 77.50 Matches: 31
Percent Similarity: 39.73% Conservative: 27
Best Local Similarity: 21.23% Mismatches: 65
Query Match: 9.15% Indels: 23
DB: 3 Gaps: 4

US-10-082-502-21 (1-162) x US-08-961-083-159 (1-1924)
OY 15 ValAspGIuLeuGIuTPrGluIleAla-----ArgValAspProLysThr 30
Db 575 ATTTCTACTTGGACAGAAAGTTGCTACTGCTCAACATCAAGTATGATTTGAAAAA 634
OY 31 IleGlnMetGlySerPheArgIleAsnProAspGlySerGlnSerValGIuValPro 50
Db 635 CTCTGTGCTGTGGC-----GATCCTGATGATGCGACAGAAAGTATGAAAGCTAAA 685
OY 51 TyRAlaArgSerGIuAlaHisLeuThrGIuLeuLeuGIuGIuValCysAspArgMetLys 70
Db .686 TTAATAAAGAGAGAGCTGAGCTAAACGCTAAACAGCTGAGTTAGCAAAAAACAAACA 745
OY 71 GIuTyRGIuGIuGlnIleAspProSerThrHisArgLysAsnTyRValArgValAspSer 90
Db 746 GAACCTGAATAAATCTTGGACAGCCTTGATCTGAAGGTAAAGCTAGAGTAATAGAT 805
OY 91 ArgAsnGlyGlySerSerGIuLeuAspLeuGlnGlyIleArgIleAspSerAspIleSer 110
Db 806 AAAGAGAGAGAAAGAGCTGATTTGATTAATAAAGCTGATGAATCAAAATAAAGTTGCT 865
OY 111 GlyThrLeuLysPheAlaCySGLuSerIleValGIuGIuTyRGIuAspGIuLeuIleGlu 130
Db 866 GATTATA-----GAAAAAGAAATTAAGTAACCTTGAA 895
OY 131 PhePheSerArgGIuAlaAspAsn-----ValLysAspLysLeuCyS 144
Db 896 ATATTACTTGGAGGGCTGATTTCTGAAGATGATAGTACTGCTGCTTCAAAATTAATTAAGCT 955
```

DB 1492 AAAAAAATTAATCAC-----ACAAAAGAACAAAATAAAAAATTAATTGAA 1536  
OY 123 GltTyGlu-----AspGluLeuIleGlu---PhePheSerArgGlu 135  
DB 1537 GATTATGAAGATGCAAAAAAGCATTTATGAAATTAATTGAAATGAAA 1596  
OY 136 AlaAspAsn-----ValIysAspLysLeuCysSerLysArg 147  
DB 1597 TTTAATAATAATTTTGACAAAGATGTCGTAGATAAATATTCAGTCGCAAGA 1647  
RESULT 5  
US-08-588-985-1  
; Sequence 1, Application US/08588985  
; Patent No. 5777094  
; GENERAL INFORMATION:  
; APPLICANT: MICHIOYUKI MATSUDA et al.  
; TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/588,985  
; FILING DATE: January 19, 1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6519 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOHETICAL:  
; ANTI-SENSE:  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE: spleen cell of homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 24..5619  
; US-08-588-985-1  
Alignment Scores:  
Pred. No.: 13.1 Length: 6519  
Score: 78.00 Matches: 44  
Percent Similarity: 40.11% Conservative: 29  
Best Local Similarity: 24.18% Mismatches: 66  
Query Match: 9.21% Indels: 44  
DB: 1 Gaps: 8  
US-10-082-502-21 (1-162) x US-08-588-985-1 (1-6519)  
OY 6 LeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluIuTrpGluIleAlaArgVal 25

DB 4660 TTACAGACCGCGTCCCTGCAGACACCCCTGAGGCCCATGAAA-----AGATCG 4707  
OY 26 AspProLysLysThrIleGlnMetLysPheArgIleAspProAspLysSerGlnSer 45  
DB 4708 AGAAGCTCAAGGACGTGATGCTTCCATGCTGCTTTTGTGGCCGAAGGATCAGATCC 4767  
OY 46 ValValGluValProGluAlaArgSerGluAlaIleLeuThrGluLeuGluGluVal 65  
DB 4768 ATG-----GAGACAAA-GTCCAGGAGGACATGAGGCCGCTC 4802  
OY 66 CysAspArgMet-----LysGluTyrgly 73  
DB 4803 CACGACAGAGATGAGGCGCTTCTTCAAAACAGTCGAGAAAAGGTGAGAAAGATGACGGC 4862  
OY 74 GluGlnIleAspProSer-----ThrIleArgLysAsnIleVal 86  
DB 4863 GTCCGATCATGCCCTCAAGCTGATGATGATGAGAGAGCCGCCGCCCTCCATGCTG 4922  
OY 87 ArgValValSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyLe---ArgIle 105  
DB 4923 CGGTCTCTCAGATGCGCTTCCATCCCGGCCCTGTGTGGCCCTGTCTTCCCTC 4982  
OY 106 AspSerAspIleSer-----GlyThrLeuLysPheAlaCysGluSerIleVal 121  
DB 4983 TCATCGACACGACACCCCTCCAGACGAGCTCCGAGGCTTCCCTCGAGCTTCCCTG 5042  
OY 122 GluGluTyrgLysArgLeuIleGluPhePheSerArgLysAlaAspArgVal---Lys 140  
DB 5043 CCAAG-----AAAATGCACTCCAGGTCCCGACGACAAAGTGCAGC 5084  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
DB 5085 GATGACCTGGAGAAAGGAGAAAGAACACAGACAGAAAGAAAAGAAACACAAACATCAA 5144  
OY 161 GluLeu 162  
DB 5145 GAGATA 5150  
RESULT 6  
US-08-971-988-1  
; Sequence 1, Application US/08971988  
; Patent No. 5786461  
; GENERAL INFORMATION:  
; APPLICANT: MICHIOYUKI MATSUDA et al.  
; TITLE OF INVENTION: CDNA OF DOCK180 GENE AND DOCK180 PROTEIN  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/971,988  
; FILING DATE: 17-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/588,985  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:

	GENERAL INFORMATION:				
	APPLICANT:	Matson, James D			
	APPLICANT:	Murison, James G.			
	TITLE OF INVENTION:	Polynucleotides, polypeptides expressed			
	TITLE OF INVENTION:	by the polynucleotides and methods for their use.			
	FILE REFERENCE:	11000.105001			
	CURRENT APPLICATION NUMBER:	US/09/724,864			
	PRIOR FILING DATE:	2000-11-28			
	PRIOR APPLICATION NUMBER:	U.S. No. 6380362 60/171,678			
	PRIOR FILING DATE:	1999-12-23			
	NUMBER OF SEQ ID NOS:	72			
	SOFTWARE:	FastSeq for Windows Version 4.0			
	SEQ ID NO 14				
	LENGTH:	915			
	TYPE:	DNA			
	ORGANISM:	Mouse			
	US-09-724-864-14				
	Alignment Scores:				
	Pred. No.:	0.00327			
	Score:	96.50			
	Percent Similarity:	42.68%			
	Best Local Similarity:	24.39%			
	Query Match:	11.39%			
	DB:	4			
	gaps:	8			
US-10-082-502-21 (1-162) x US-09-724-864-14 (1-915)					
OY	6	LenuHSCysGLyAlaCysArgAlaLeuValAspGluLeuGluTrpGluIleAlaArgVal	25		
Db	158	CTGGCGTCACAGTCCGTGCCTGCCGCTCCAGATGGGGCAAGCTTGCGAAGAACA	217		
OY	26	AspProLysLysThrIleGlnMetClySerPheArgIleAsnProAspGlySerGlnSer	45		
Db	218	GAGCGTAAATCTCAC-----ACTCCAGAGCCGACG---GGA	250		
OY	46	ValValGluValIPROYrAlaArgSerGluAlaHisIleuThrGluLeuGluGluVal	65		
Db	251	TTWGAGGAGGCTG-----AGTGAATCCACGTAACAGATGGCTCGACACAGCC	298		
OY	66	CysAspArg--MettLysGluTrpGlyGluGlnIleAspProSerThrHisArgLysAsn	84		
Db	299	TGCTCTCAGAACGTGGCAGTCTCATGA-----GTTCAATGAAGTAGAC	340		
OY	85	TyrValArgValAlaSerArgAsnGlyGluSerSerGluLeuAspLeuGlnGlyIleArg	104		
Db	341	CAGATGAAGCGCTCTCACGGGCCACGACTTAAGCAAAGGGCCAGAG-----CCAAGA	391		
OY	105	IleAspSerAspIleSerGly-----ThreuleylPheAlaCysGluSer	119		
Db	392	ATCACCGCGATGATTTCGTGGGGGCCCTCGGCCCAATAGAGCTCTCCAGACAGCTGTTCCAC	451		
OY	120	IleValGluGluTrp--GluAspGluLeuIleGluPheSerArgGluAlaAspAsn	138		
Db	452	TACCTGGGTGAGTTTGGAGAGAGACCAATCATATGATAACCTTACCGCCAAAGCCAAAGCAAT	511		
OY	139	ValLysAspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaIleuHisArgSer	158		
Db	512	CTGGAGCGCGCTCTCTGTGGGGGCACCACCAATGGGCCCTGTCTACAGGAGATCTGGCCACG	571		
OY	159	HisAspGluLeu	162		
Db	572	AGAGAGAGCTT	583		
RESULT 4					
US-08-257-073-10					
; Sequence 10, Application US/08257073					
; Patent No. 576597					
; GENERAL INFORMATION:					
; APPLICANT:	Paoletti, Enzo				
; APPLICANT:	de Taisne, Charles				
; APPLICANT:	Tine, John A.				
; TITLE OF INVENTION:	MALARIA RECOMBINANT POXYVIRUS VACCINE				

```

NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSER: Curtiss, Morris & Safford, P.C.
STREET: 530 Fifth Avenue, 25th Floor
CITY: New York
STATE: New York
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/257,073
FILING DATE: 09-JUN-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,783
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/852,305
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/672,183
FILING DATE: 20-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2570
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTIS
INFORMATION FOR SEO ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 5181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-257-073-10

Alignment Scores:
Pred. No.: 0.654 Length: 5181
Score: 87.00 Matches: 36
Percent Similarity: 43.95% Conservative: 33
Best Local Similarity: 22.93% Mismatches: 58
Query Match: 10.27% Indels: 30
DB: 1 Gaps: 7

US-10-082-502-21 (1-162) x US-08-257-073-10 (1-5181)
QY 16 AspGluLeuGIUrrpGiuIleAlaAaGValAspProLyLysThrIleGlnMetGlySer 35
||| ||||| ||| :|: ||| :|: :|:
Db 1192 GATCACTTGAATTAGAATTATTTTAAGAGAAAAAAATAAAGTAGTGATACACT 1251
QY 36 PheArgIleAsnProAspGlySerGlnSerValGlValAlProTyAlaArgSerGlu 55
:::||| | ||| :|: :|: |||||
Db 1252 AAATCACACMAATGCCTACGAATCTGTTCACAATACCAAAGTCCCTAACCAATGGTATT 1311
QY 56 AlaHis-----LeuThrGluLeuLeuGluGluValCys-----AspArgMetLys 70
:|: :|: ||||| :|: :|: :|:
Db 1312 GTAATATCTTTACCACCTACCTCGATTTATCATTAATTCATTAGCTGCAGATTAATGATAAAT 1371
QY 71 GluTyrgIleGluGlnIleAspProSerThrHisArgLysAsnTyValArgValAlser 90
|||||:|: :|: ||| ||| :|: :|: :|:
Db 1372 TCATATGGTATTGAATGAATCTGATACTAAAGAAAAAATTAAGAAAATTTTAC 1431
QY 91 ArgAsnIeGlu-----SerSerGluLeuAspLeuGlnGly 102
||| ||| :|: :|: |||||:|:
Db 1432 GATTAATAGAGAAAATAATTCATTAAATTAACATTAAAAAAAACAATGATTAGAACAA 1491
QY 103 IleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIleValGlu 122

```





GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 30, 2003, 22:33:58 ; Search time 35.3198 Seconds  
(without alignments)  
1406.624 Million cell updates/sec

Title: US-10-082-502-21  
Perfect score: 847  
Sequence: 1 RRSQDHRGACRALVDELEW.....LCSKRIDLCDHALHRSDEL 162

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlj  
-O=/cg2\_1/USPTO.spool/US10082502/rnat\_24032003\_135100\_6616/app-query.fasta.1.1308  
-DB=Issued\_Patents\_NA -OEM=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsun62 -TRANS=human40.ccd  
-LIST=45 -DOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US10082502\_6CGN\_1.131.0rnat\_24032003\_135100\_6616 -NCPD=6 -ICPD=3  
-NO\_XLPAY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -NAT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -YGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

Issued\_Patents\_NA: \*  
1: /cg2\_6/ptodata/1/lna/5A.COMB.seq: \*  
2: /cg2\_6/ptodata/1/lna/5B.COMB.seq: \*  
3: /cg2\_6/ptodata/1/lna/6A.COMB.seq: \*  
4: /cg2\_6/ptodata/1/lna/6B.COMB.seq: \*  
5: /cg2\_6/ptodata/1/lna/PCU05.COMB.seq: \*  
6: /cg2\_6/ptodata/1/lna/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	132	15.6	744	4	US-09-163-285-3
2	132	15.6	1512	4	US-09-163-285-1
3	96.5	11.4	915	4	US-09-724-864-14
4	87	10.3	5181	1	US-08-257-073-10
5	78	9.2	6519	1	US-08-588-985-1
6	78	9.2	6519	1	US-08-971-988-1
7	77.5	9.1	1924	3	US-08-961-083-159
8	76.5	9.0	3752	4	US-08-961-527-208
9	76	9.0	4131	1	US-07-872-644-38
10	76	9.0	4131	1	US-08-297-494-38
11	76	9.0	4131	1	US-08-297-510-38
12	76	9.0	4131	1	US-08-479-532-38

13	76	9.0	4131	1	US-08-455-526-38	Sequence 38, Appl
14	76	9.0	4131	1	US-08-455-525-38	Sequence 38, Appl
15	76	9.0	4131	3	US-09-139-491-38	Sequence 38, Appl
16	76	9.0	4131	5	PCT-US92-03322-38	Sequence 38, Appl
17	76	9.0	5558	4	US-08-961-527-103	Sequence 103, App
18	75	8.9	5661	4	US-08-938-105-2	Sequence 2, Appl
19	74.5	8.8	1216	4	US-09-071-035-3	Sequence 3, Appl
20	74.5	8.8	1347	4	US-09-071-035-1	Sequence 1, Appl
21	74.5	8.8	4371	1	US-08-803-973-1	Sequence 1, Appl
22	74.5	8.8	4371	1	US-08-803-972-1	Sequence 9, Appl
23	74	8.7	918	3	US-08-676-444-1	Sequence 41, Appl
24	74	8.7	1640	4	US-08-676-444-1	Sequence 42, Appl
25	74	8.7	3789	1	US-07-872-644-42	Sequence 42, Appl
26	74	8.7	3789	1	US-08-297-494-42	Sequence 42, Appl
27	74	8.7	3789	1	US-08-297-510-42	Sequence 42, Appl
28	74	8.7	3789	1	US-08-479-532-42	Sequence 42, Appl
29	74	8.7	3789	1	US-08-455-526-42	Sequence 42, Appl
30	74	8.7	3789	1	US-08-455-525-42	Sequence 42, Appl
31	74	8.7	3789	3	US-09-139-491-42	Sequence 42, Appl
32	74	8.7	3789	5	PCT-US92-03322-42	Sequence 42, Appl
33	73.5	8.7	633	4	US-09-134-001C-578	Sequence 578, App
34	73.5	8.7	833	2	US-08-837-029-1	Sequence 1, Appl
35	73.5	8.7	4865	3	US-08-894-017-24	Sequence 24, Appl
36	73.5	8.7	5177	6	5352450-1	Patent No. 5352450
37	73.5	8.7	12127	4	US-08-961-527-148	Sequence 148, App
38	72.5	8.6	41708	4	US-09-470-512A-3	Sequence 3, Appl
39	72	8.5	640	3	US-08-961-083-33	Sequence 33, Appl
40	72	8.5	11309	4	US-08-961-527-108	Sequence 108, App
41	71.5	8.4	2088	4	US-09-458-481B-3	Sequence 3, Appl
42	71.5	8.4	4394	2	US-08-750-152A-1	Sequence 1, Appl
43	71.5	8.4	5084	1	US-08-306-691B-21	Sequence 21, Appl
44	71.5	8.4	5084	5	PCT-US93-06251-25	Sequence 25, Appl
45	71	8.4	601	4	US-09-556-877-42	Sequence 22, Appl

#### ALIGNMENTS

RESULT 1  
Sequence 3, Application US/09163285  
Patent No. 6204013  
GENERAL INFORMATION:  
APPLICANT: Rhodadoust, Mehran  
TITLE OF INVENTION: NOVEL MSP-5 PROTEIN AND NUCLEIC ACID MOLECULES  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,285  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/090,398  
FILING DATE: June 24, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: MNI-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:

10

11

12

```
Db 603 CGAGGAGCCAGAGTCTCCAGTGTGAGACAGAGGCTCTGTGTGATGAATGAG 662
Qy 21 GUILLEALARGVALASPProLYSThrILEGIMeGLYSerPHEArgILEAsnPro 40
Db 663 GAAATTGGCCAGGTGGACCCCAAGAGACATTCAGATGGGATCTTCCGGATCAATCCA 722
Qy 41 ASPGLYSerGLINSerVALIGLUVALProTYRALAArgSERGLUALAHISLeuThrGLU 60
Db 723 GATGGCACCCAGTCAGTGTGGAGTGGCTTATGCCGCTCAGAGGCCACCCTCACAGAG 782
Qy 61 LeuLeuGLUGLUVALCYAsPARMeTLYSGIuTYRGLYGLUGLINILEAsnProSerThr 80
Db 783 CTGCTGGAGAGATATGTGACCGGATGAGAGGATGGGAGACAGATTGATCTTCCACC 842
Qy 81 HISArgLYSAsnTYRVALArgVALISerARGAsnGLYGLUSerSERGLULeuAspLeu 100
Db 843 CATCCAGAACTAGTACGTACGTAGTGGCCGGATGGAGATCGAATCCAGTAAGTGAACCTA 902
Qy 101 GINGLYILEArgILEAspSERAspILESerGLYThrLeuLYSPHEALACYSGLUSerILE 120
Db 903 CAAGGCATCCGATCGACTCAGATATTAGCGGCACCTCAAGTTGCGTGTGAGAGCAT 962
Qy 121 VALGLUGIuTYRGLUAspGLUleuTLEGLIuPHESerARGLUALAspAsnVALYS 140
Db 963 GTGGAGCAATACGAGGATGAATCAATTGAATCTTTCCCGAGAGCTGACAAATGTAAA 1022
Qy 141 AspLYSLeuCYSerLYSArgTHRAspLeuCYAspHISALeUHLISArgSerHISAsp 160
Db 1023 GACAAACCTTTCAGTAAAGCAACATCTTGTGACCATGCCCTGCACATATCGCATGAT 1082
Qy 161 GLUleu 162
Db 1083 GAGCTA 1088
```

Search completed: March 31, 2003, 11:59:14  
Job time : 51.9767 secs

```
|||||
DB 843 CATCGCAGAACTACTACTGCTAGTGGCGCGGAATGGAAATCCAGTGAACGGACCTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCTCCGAAATCGACTCGATGATATTAGGGACCCCTCAAGTTGGTGTGAGACATT 962
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
DB 963 GTGGAGGAATACGAGATGAACTGATGAACTCTTTCCGAGAGGCTGACAAATGTTAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAACTTTCAGTAAAGCAGACAGATCTTTGTGACCATGCCCCCTGCACATATCGATGAT 1082
OY 161 GluLeu 162
DB 1083 GAGCTA 1088

RESULT 14
US-10-142-419-173
; Sequence 173, Application US/10142419
; Publication No. US20030044945A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C244
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-142-419-173

Alignment Scores:
Pred. No.: 3,41e-99 Length: 1210
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-21 (1-162) x US-10-142-419-173 (1-1210)
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
DB 603 CGGAGGAGCCAGATCTCCACTGTGGAGCATGCGGCGCTGTGGTGAATGAACTGAAATGG 662
OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAsnPro 40
DB 663 GAAATTCGCCAGGTGAGACCCCAAGACAAATTCAGATGGGATCTTCCGGAATCAATCA 722
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
```

```
|||||
DB 723 GATGAGCCAGCAGTACTGGTGGAGTGCCCTTATATGCCGCTCAGAGGCCACACAGAG 782
OY 61 LeuLeuGluGluValLysAspArgMetLysGluTyrGlyGluGlnIleAspProSerTrp 80
DB 783 CTCTGTGAGAGATATGATGACCGGATGAAAGAGTATGGGAACAGATTAATCTTCACAC 842
OY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyGluSerSerGluLeuAspLeu 100
DB 843 CATCGCAGAACTACTACTGCTAGTGGCGCGGAATGGAAATCCAGTGAACGGACCTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCTCCGAAATCGACTCGATGATATTAGGGACCCCTCAAGTTGGTGTGAGACATT 962
OY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
DB 963 GTGGAGGAATACGAGATGAACTGATGAACTCTTTCCGAGAGGCTGACAAATGTTAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAACTTTCAGTAAAGCAGACAGATCTTTGTGACCATGCCCCCTGCACATATCGATGAT 1082
OY 161 GluLeu 162
DB 1083 GAGCTA 1088

RESULT 15
US-10-123-262-173
; Sequence 173, Application US/10123262
; Publication No. US20030049816A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C38
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-123-262-173

Alignment Scores:
Pred. No.: 3,41e-99 Length: 1210
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-21 (1-162) x US-10-123-262-173 (1-1210)
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20
```

|||||  
DB 1083 GAGCTA 1088

## RESULT 12

US-10-143-114-173

; Sequence 173, Application US/10143114  
; Publication No. US20030036180A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C211  
; CURRENT FILING DATE: 2002-05-09  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien

US-10-143-114-173

## Alignment Scores:

Pred. No.: 3,41e-99 Length: 1210  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: Gaps: 0

US-10-082-502-21 (1-162) x US-10-143-114-173 (1-1210)

OY 1 ArgatgSerGlnAspLeuHicysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGAGAGGAGCGAGATCTCCAGTGTGAGCAGTGCAGGGCTGTGTGAGTAAGTGG 662  
OY 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 663 GAATATGCCAGGTGGAGCCCAAGAACATTCAGATGGGATCTTCCGGATCAATCA 722  
OY 41 AspGlySerGlnSerValAlaGluValProTyraAlaArgSerGluAlaHisLeuThrGlu 60  
DB 723 GATGACACCACTCAGTGTGAGGTGCTTATGCCCCCTCAGAGCCCACTCACAAG 782  
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80  
DB 783 CTGCTGGAGGAGATGTGACCGGATGAAGAGTATGGGGAACAGATTGATCTTCCACC 842  
OY 81 HisArgLysAsnTyraValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 100  
DB 843 CATCCCAAGAACTAGTACGTGTAGTGGCCGGAATGCAGAAATCCAGTGAATGAGACTA 902  
OY 101 GlnGlyIleArgGlyLeaAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 903 CAAGGCATCCGAAATGATGAGATGTAGCGGACCACTCAAGTTGGCTGTGAGAGCATT 962  
OY 121 ValGluGluTyrglyGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140

|||||  
DB 963 GTGAGGAATACAGAGATGAACTCATGTGATCTTTCCCGAGAGCGTGAATTTAAA 1022

## OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspPheAlaLeuHisArgSerHisAsp 160

DB 1023 GACAAACTTTGACAGTACGAGAACAGATCTTTGTGACATGCCCTGCACATATCGCATGAT 1082

## OY 161 GluLeu 162

DB 1083 GAGCTA 1088

## RESULT 13

US-10-140-002-173

; Sequence 173, Application US/10140002  
; Publication No. US20030037623A1

## GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C59  
; CURRENT APPLICATION NUMBER: US/10/140,002  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 173  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: Homo Sapien

US-10-140-002-173

## Alignment Scores:

Pred. No.: 3,41e-99 Length: 1210  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: Gaps: 0

US-10-082-502-21 (1-162) x US-10-140-002-173 (1-1210)

OY 1 ArgatgSerGlnAspLeuHicysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGAGAGGAGCGAGATCTCCAGTGTGAGCAGTGCAGGGCTGTGTGAGTAAGTGG 662  
OY 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 663 GAATATGCCAGGTGGAGCCCAAGAACATTCAGATGGGATCTTCCGGATCAATCA 722  
OY 41 AspGlySerGlnSerValAlaGluValProTyraAlaArgSerGluAlaHisLeuThrGlu 60  
DB 723 GATGACACCACTCAGTGTGAGGTGCTTATGCCCCCTCAGAGCCCACTCACAAG 782  
OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrglyGluGlnIleAspProSerThr 80  
DB 783 CTGCTGGAGGAGATGTGACCGGATGAAGAGTATGGGGAACAGATTGATCTTCCACC 842  
OY 81 HisArgLysAsnTyraValArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 100

```
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C162
CURRENT APPLICATION NUMBER: US/10/140,474
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 173
LENGTH: 1210
TYPE: DNA
ORGANISM: Homo Sapien
US-10-140-474-173
```

```
Alignment Scores:
Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%   Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                    Gaps: 0
```

US-10-082-502-21 (1-162) x US-10-140-474-173 (1-1210)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGAGAGGAGCCAGATCTCCAGCTGTGGAGCATGCGAGGCTGTGGTGGATGACATGAGATGG 662
OY 21 GlnIleAlaArgValAspProLysIleGlnMetCysPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGAGCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCCA 722
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60
DB 723 GATGGCAGCCAGATCGATGAGGAGTCCCTTATGCCCGCTCAGAGGCCCTCAGACAGAG 782
OY 61 LeuLeuGlnGluValCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80
DB 783 CTCTGAGAGAGATATGTGTGACCGGATGAGAGATGCGGAGATGATCTTCACACC 842
OY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTACGATCTGTAGTGGCCGGAATGGAAATCCATGAACTGGACCTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCTCCGAAATCCGATCATGATATTAGCGCACCTCAAGTTTGGCTGTGAGAGCAT 962
OY 121 ValGlnGluTyrGlnAspGluLeuIleGluPheSerArgGlnAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGATGATGATCATGATTTCTTTCCGAGAGGCTGCAATGTAAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAAACTTTGACAGTAAAGCAACAGATCTTGTGACCATCCCTGACAAATATGCGAT 1082
OY 161 GluLeu 162
DB 1083 GAGCTA 1088
```

RESULT 11  
US-10-142-431-173  
Sequence 173, Application US/10142431  
Publication No. US20030036179A1  
GENERAL INFORMATION:

```
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Flvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C251
CURRENT APPLICATION NUMBER: US/10/142,431
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 173
LENGTH: 1210
TYPE: DNA
ORGANISM: Homo Sapien
US-10-142-431-173
```

```
Alignment Scores:
Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%   Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                    Gaps: 0
```

US-10-082-502-21 (1-162) x US-10-142-431-173 (1-1210)

```
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
DB 603 CGAGAGGAGCCAGATCTCCAGCTGTGGAGCATGCGAGGCTGTGGTGGATGACATGAGATGG 662
OY 21 GlnIleAlaArgValAspProLysIleGlnMetCysPheArgIleAsnPro 40
DB 663 GAATATGCCAGGTGGAGCCCAAGAGACCATTCAGATGGATCTTCCGATCAATCCA 722
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60
DB 723 GATGGCAGCCAGATCGATGAGGAGTCCCTTATGCCCGCTCAGAGGCCCTCAGACAGAG 782
OY 61 LeuLeuGlnGluValCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80
DB 783 CTCTGAGAGAGATATGTGTGACCGGATGAGAGATGCGGAGATGATCTTCACACC 842
OY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
DB 843 CATCGCAAGAACTACGATCTGTAGTGGCCGGAATGGAAATCCATGAACTGGACCTA 902
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
DB 903 CAAGGCTCCGAAATCCGATCATGATATTAGCGCACCTCAAGTTTGGCTGTGAGAGCAT 962
OY 121 ValGlnGluTyrGlnAspGluLeuIleGluPheSerArgGlnAlaAspAsnValLys 140
DB 963 GTGAGAGAAATACAGATGATGATCATGATTTCTTTCCGAGAGGCTGCAATGTAAA 1022
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
DB 1023 GACAAACTTTGACAGTAAAGCAACAGATCTTGTGACCATCCCTGACAAATATGCGAT 1082
OY 161 GluLeu 162
```

```

: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-176-921-173

```

## Alignment Scores:

```

Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%    Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                      Gaps: 0

```

US-10-082-502-21 (1-162) x US-10-176-921-173 (1-1210)

```

OY 1  ARGATGSRGINASPLEUHNISCYSGLYALACYSATGALALEUVALASPGLULEUGLUTRP 20
    |||||||
DB 603 CGGAGGACCCAGGATCTCCACGTGGAGCATGSCAGGCTCTGGATGAACTAGCAATGG 662
OY 21  GLUIEALAAAGVALASPRGOLYSLSYTHRIIEGIMETGLYSERPHARGILLEASNP 40
    |||||||
DB 663 GAAATGGCCAGGTGGACCCCAAGAAAGCAATTCAGATGGAGTCTTCCGGATCAATCCA 722
OY 41  ASPLGYSERGINSERVALVALGLUVALPROTYRALAARGSERGLUALAHISLEUTHRG 60
    |||||||
DB 723 GATGGCACCCAGTCAAGTGTGAGGTGAGTATGCGCTATGCGCCCTCAGAGGCCACCTCAGAG 782
OY 61  LEULEUGLUGLUALCYASAPRGMETLYSGIUTYRGYUGLUGLILAEAPPROSETHR 80
    |||||||
DB 783 CTGCTGGAGGAGATATGTGACCGGATGGAAGGATGGGGAACAGATTGATCTCCACC 842
OY 81  HISARGLYSANTYRVALARVALVALSERARGASNGLYGUSERSERGLULEUASPLEU 100
    |||||||
DB 843 CATCCCAAGAACTAGCTAGTGTACTGGCGCGGAGATGAGATCCAGTGAACCTGACCTA 902
OY 101  GINGLYILEARGILEASPSERAPRIIESERGILYTHLEULYRPHALACYSGLUSERL 120
    |||||||
DB 903 CAAGCACCACCGAATGCACTCAAGATATTAGCGCACCCCTCAAGTTGGCTGTGAGACATT 962
OY 121  VALGUGLUTYRGUASPGLULEULIEGLUPHERPESERARGLUALAASPAASNYALLYS 140
    |||||||
DB 963 GTGGAGGAATAGAGAGATGAACTCATTTGAATCTTTCCGAGAGGCGGACATGTTAAA 1022
OY 141  ASPLYSLEUCYSSELYSARGTTHRASPLEUCYSASPHISALALEUHNISARGSERHIS 160
    |||||||
DB 1023 GACAAACTTTCAGTAAAGCAAGACAGATCTTTGTGACATGCCCTGCACATATCCGATGAT 1082
OY 161  GLULEU 162
    |||||||
DB 1083 GAGCTA 1088

```

## RESULT 9

```

: Sequence 173, Application US/10137865
: Publication No. US20030032155A1

```

## GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Collin K
: APPLICANT: Wood, William

```

```

: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C154
: CURRENT APPLICATION NUMBER: US/10/137, 865
: PRIOR APPLICATION REMOVED - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 173
: LENGTH: 1210
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-137-865-173

```

## Alignment Scores:

```

Pred. No.: 3,41e-99      Length: 1210
Score: 830.00           Matches: 158
Percent Similarity: 98.77%      Conservative: 2
Best Local Similarity: 97.53%    Mismatches: 2
Query Match: 97.99%           Indels: 0
DB: 9                      Gaps: 0

```

US-10-082-502-21 (1-162) x US-10-137-865-173 (1-1210)

```

OY 1  ARGATGSRGINASPLEUHNISCYSGLYALACYSATGALALEUVALASPGLULEUGLUTRP 20
    |||||||
DB 603 CGGAGGACCCAGGATCTCCACGTGGAGCATGSCAGGCTCTGGATGAACTAGCAATGG 662
OY 21  GLUIEALAAAGVALASPRGOLYSLSYTHRIIEGIMETGLYSERPHARGILLEASNP 40
    |||||||
DB 663 GAAATGGCCAGGTGGACCCCAAGAAAGCAATTCAGATGGAGTCTTCCGGATCAATCCA 722
OY 41  ASPLGYSERGINSERVALVALGLUVALPROTYRALAARGSERGLUALAHISLEUTHRG 60
    |||||||
DB 723 GATGGCACCCAGTCAAGTGTGAGGTGAGTATGCGCTATGCGCCCTCAGAGGCCACCTCAGAG 782
OY 61  LEULEUGLUGLUALCYASAPRGMETLYSGIUTYRGYUGLUGLILAEAPPROSETHR 80
    |||||||
DB 783 CTGCTGGAGGAGATATGTGACCGGATGGAAGGATGGGGAACAGATTGATCTCCACC 842
OY 81  HISARGLYSANTYRVALARVALVALSERARGASNGLYGUSERSERGLULEUASPLEU 100
    |||||||
DB 843 CATCCCAAGAACTAGCTAGTGTACTGGCGCGGAGATGAGATCCAGTGAACCTGACCTA 902
OY 101  GINGLYILEARGILEASPSERAPRIIESERGILYTHLEULYRPHALACYSGLUSERL 120
    |||||||
DB 903 CAAGCACCACCGAATGCACTCAAGATATTAGCGCACCCCTCAAGTTGGCTGTGAGACATT 962
OY 121  VALGUGLUTYRGUASPGLULEULIEGLUPHERPESERARGLUALAASPAASNYALLYS 140
    |||||||
DB 963 GTGGAGGAATAGAGAGATGAACTCATTTGAATCTTTCCGAGAGGCGGACATGTTAAA 1022
OY 141  ASPLYSLEUCYSSELYSARGTTHRASPLEUCYSASPHISALALEUHNISARGSERHIS 160
    |||||||
DB 1023 GACAAACTTTCAGTAAAGCAAGACAGATCTTTGTGACATGCCCTGCACATATCCGATGAT 1082
OY 161  GLULEU 162
    |||||||
DB 1083 GAGCTA 1088

```

## RESULT 10

```

: Sequence 173, Application US/10140474
: Publication No. US20030032156A1

```

## GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey

```



```
QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
    |||||
Db 603 CGAGAGACCGACGATCTCCACTGTGAGCAGATCAGGGCTCTGCTGATCAACTAGCAATGG 662
QY 21 GluIleAlaArgValAspProLysTrpHisLeuMetCysSerPheArgIleAsnPro 40
    |||||
Db 663 GAATATGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuTrpGlu 60
    |||||
Db 723 GATGGACGCGACGATCTGCTGAGAGTGGCTTATAGCCCGCTCAAGGCCCACTCAGAGAG 782
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||||
Db 783 CTCCTGAGAGAGATATGTGACCGCATCAAGAGCATATGGCGAACAAGATTCCTTCACC 842
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
    |||||
Db 843 CATCGCAAGAACTACGATCTGATGGGCGCAATGGAGAAATCCAGTAAGTGGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||
Db 903 CAAGGACATCGAATCGACTCAGATATTAGGGCAACCTCAAGTTTGGCTGAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||||
Db 963 GTGAGAGAAATACAGAGATCACTCATTTGATCTTCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||||
Db 1023 GACAAATCTTGACGATTAAGCAACAGATCTTGTGACCATCCCTGCACATATGCGCATGAT 1082
QY 161 GluLeu 162
    |||||
Db 1083 GAGCTA 1088

RESULT 7
US-10-176-918-173
; Sequence 173, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-918-173

Alignment Scores: 3.41e-99 Length: 1210
Pred. No.:
```

```
Score: 830.00 Matches: 158
Percent Similarity: 98.77% Conservative: 2
Best Local Similarity: 97.53% Mismatches: 2
Query Match: 97.99% Indels: 0
DB: 9 Gaps: 0

US-10-082-502-21 (1-162) x US-10-176-918-173 (1-1210)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTrp 20
    |||||
Db 603 CGAGAGACCGACGATCTCCACTGTGAGCAGATCAGGGCTCTGCTGATCAACTAGCAATGG 662
QY 21 GluIleAlaArgValAspProLysTrpHisLeuMetCysSerPheArgIleAsnPro 40
    |||||
Db 663 GAATATGCCAGGTGGAGCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 722
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuTrpGlu 60
    |||||
Db 723 GATGGACGCGACGATCTGCTGAGAGTGGCTTATAGCCCGCTCAAGGCCCACTCAGAGAG 782
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||||
Db 783 CTCCTGAGAGAGATATGTGACCGCATCAAGAGCATATGGCGAACAAGATTCCTTCACC 842
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
    |||||
Db 843 CATCGCAAGAACTACGATCTGATGGGCGCAATGGAGAAATCCAGTAAGTGGACCTA 902
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||
Db 903 CAAGGACATCGAATCGACTCAGATATTAGGGCAACCTCAAGTTTGGCTGAGAGCAT 962
QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
    |||||
Db 963 GTGAGAGAAATACAGAGATCACTCATTTGATCTTCCGAGAGGCTGACAAATGTTAAA 1022
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||||
Db 1023 GACAAATCTTGACGATTAAGCAACAGATCTTGTGACCATCCCTGCACATATGCGCATGAT 1082
QY 161 GluLeu 162
    |||||
Db 1083 GAGCTA 1088

RESULT 8
US-10-176-921-173
; Sequence 173, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Laureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
```





PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/059278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/059334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/059694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086414  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086430  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088730  
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088741  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

Alignment Scores:  
Pred. No.: 3,41e-99 Length: 1210  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2  
Query Match: 97.99% Indels: 0  
DB: 9 Gaps: 0

US-10-082-502-21 (1-162) x US-10-028-072-173 (1-1210)

QY 1 ArgArGserGlnAspPleuHIsCyGlyAlaCyArGAlaLeuValAspGluLeuGluTrp 20  
DB 603 CGGAGGAGCCGAGATCTCCACTGTGGAGCAGGCGCTGTGGATGAACCTGAATGG 662  
QY 21 GluIleAlaArGValAspProLySThrIleGlnMetGlySerPheArgIleAspPro 40  
DB 663 GAAATTCGCCAGGTGGAGCCCAAGAACACCTTACATGGGATCTTCCGATCAATCCA 722  
QY 41 AspGlySerGlnSerValIleGluValProTyrlaArGserGluAlaHIsLeuThrGlu 60  
DB 723 GATGGCAGCCAGTCAGTGTGGAGTGCCTTATGCCGCTCAGAGGCCACCTCAGAG 782  
QY 61 LeuLeuGluGluValCysAspArgMetIleGluTyrlGluGlnIleAspProSerThr 80  
DB 783 CTGCTGGAGGAGATGTGTGACCGGATGAAGATATGGGAAAGATTGATTCCTCCACC 842  
QY 81 HisArgLySAsnTyrlValArGValIleSerArgAsnGlyIleSerSerGluLeuAspLeu 100  
DB 843 CATGCCAAGAACCTACGATCGTGTAGTGGCCGGAATGAGAAATCATGTAATGACCTA 902  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLySPhelaIaCySgluSerIle 120  
DB 903 CAAGCATCCGAATCGACTCAGATATTAGCGCACCTCAAGTTGGGTGTGACAGCATTT 962  
QY 121 ValGluGluTyrlGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLyS 140  
DB 963 GTGGAGGAATACGAGGATGAACCTCATTTCTTCCGAGAGGCTGACAAATGTAAA 1022  
QY 141 AspLySLeuCySserLySArgThrAspLeuCySAspHIsAlaLeuHIsArgSerHIsAsp 160  
DB 1023 GACAAATTTGCACTAAGCGAACAGATCTTGTGACCATGCCCTGCACATATTCGATGAT 1082



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 07:02:46 ; Search time 48.9767 Seconds  
(without alignments)  
2814.757 Million cell updates/sec

Title: US-10-082-502-21

Perfect score: 847  
Sequence: 1 RRSQDHCACACALVDELEW.....LCSSKRDLCDAHLRSHDEL 162

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 574371 segs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=framed\_p2n.model -DEV=xlp  
-O=/cg2\_1/USPTO.spool/US10082502/runtat\_24032003\_135103\_6870/app\_query.fasta.1.1308  
-DB=Published.Applications.NA -OPMT=fastap -SUFFIX=inp -MINMATCH=0.1  
-LOOPCU=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62  
-TRANS=human40.cdt -LIST=45 -DCCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MALEN=2000000000 -USER=US10082502.ecgn.1.1.145 -runtat\_24032003\_135103\_6870  
-NCPU=6 -ICPU=3 -NO\_XLIPY -NO\_MMAP -LARGEDQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV.TIMEOUT=120 -MARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published.Applications.NA.\*

- 1: /cg2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cg2\_6/ptodata/2/pubpna/PC7\_NEW\_PUB.seq.\*
- 3: /cg2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cg2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq.\*
- 5: /cg2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cg2\_6/ptodata/2/pubpna/PC7US\_PUBCOMB.seq.\*
- 7: /cg2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cg2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cg2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cg2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cg2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cg2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cg2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cg2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830	98.0	814	10	US-09-864-864-309
2	830	98.0	1210	9	US-10-028-072-173
3	830	98.0	1210	9	US-10-121-049-173
4	830	98.0	1210	9	US-10-123-904-173

5	830	98.0	1210	9	US-10-140-470-173	Sequence 173, App
6	830	98.0	1210	9	US-10-175-746-173	Sequence 173, App
7	830	98.0	1210	9	US-10-176-918-173	Sequence 173, App
8	830	98.0	1210	9	US-10-176-921-173	Sequence 173, App
9	830	98.0	1210	9	US-10-137-865-173	Sequence 173, App
10	830	98.0	1210	9	US-10-140-474-173	Sequence 173, App
11	830	98.0	1210	9	US-10-142-431-173	Sequence 173, App
12	830	98.0	1210	9	US-10-143-114-173	Sequence 173, App
13	830	98.0	1210	9	US-10-140-002-173	Sequence 173, App
14	830	98.0	1210	9	US-10-142-419-173	Sequence 173, App
15	830	98.0	1210	9	US-10-123-262-173	Sequence 173, App
16	830	98.0	1210	9	US-10-142-423-173	Sequence 173, App
17	830	98.0	1210	9	US-10-121-050-173	Sequence 173, App
18	830	98.0	1210	9	US-10-141-755-173	Sequence 173, App
19	815	96.2	657	9	US-09-822-946-198	Sequence 398, App
20	154	18.2	564	10	US-09-864-864-228	Sequence 228, App
21	132	15.6	1457	10	US-09-846-5738-110	Sequence 10, App1
22	132	15.6	1470	9	US-10-035-855-58	Sequence 58, App1
23	132	15.6	1470	9	US-10-035-855-58	Sequence 58, App1
24	132	15.6	1470	9	US-10-174-590-479	Sequence 479, App
25	132	15.6	1470	9	US-10-176-758-479	Sequence 479, App
26	132	15.6	1470	9	US-10-175-737-479	Sequence 479, App
27	132	15.6	1470	9	US-10-173-706-479	Sequence 479, App
28	132	15.6	1470	9	US-10-175-738-479	Sequence 479, App
29	132	15.6	1470	9	US-10-175-752-479	Sequence 479, App
30	132	15.6	1470	9	US-10-176-482-479	Sequence 479, App
31	132	15.6	1470	9	US-10-176-757-479	Sequence 479, App
32	132	15.6	1470	9	US-10-176-913-479	Sequence 479, App
33	132	15.6	1470	9	US-10-180-552-479	Sequence 479, App
34	132	15.6	1470	9	US-10-180-557-479	Sequence 479, App
35	132	15.6	1470	9	US-09-931-836-58	Sequence 58, App1
36	132	15.6	1470	9	US-10-173-700-479	Sequence 479, App
37	132	15.6	1470	9	US-10-174-572-479	Sequence 479, App
38	132	15.6	1470	9	US-10-174-579-479	Sequence 479, App
39	132	15.6	1470	9	US-10-174-582-479	Sequence 479, App
40	132	15.6	1470	9	US-10-174-588-479	Sequence 479, App
41	132	15.6	1470	9	US-10-175-739-479	Sequence 479, App
42	132	15.6	1470	9	US-10-175-740-479	Sequence 479, App
43	132	15.6	1470	9	US-10-175-743-479	Sequence 479, App
44	132	15.6	1470	9	US-10-176-488-479	Sequence 479, App
45	132	15.6	1470	9	US-10-176-492-479	Sequence 479, App

#### ALIGNMENTS

RESULT 1  
US-09-864-864-309  
Sequence 309, Application US/09864864  
Patent No. US20020102679A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jianshun  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Dillon, David C.  
APPLICANT: Secrist, Heather  
APPLICANT: Lodes, Michael J.  
APPLICANT: Algate, Paul A.  
APPLICANT: Flinn, Steve P.  
APPLICANT: Mannion, Jane  
APPLICANT: Benson, Darin R.  
APPLICANT: Carter, Derrick  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.523  
CURRENT APPLICATION NUMBER: US/09/864, 864  
NUMBER OF SEQ ID NOS: 341  
SOFTWARE: Corixa Invention Disclosure Database  
SEQ ID NO 309  
LENGTH: 814  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-864-864-309

Query Match: 97.99% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-21 (1-162) x BM832828 (1-588)

```

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuTyr 20
    |||||||
DB 62 CGAGAGAGCCAGATCTCCACTGTGGATCGAGGGCTGTGGATGAACTAGATGG 121
    |||||||
QY 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40
    |||||||
DB 122 GAAATGGCCAGGTGACCCCAAGAACCATTCAGATGGATCTTCCGATCAATCCA 181
    |||||||
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
    |||||||
DB 182 GATGGCAGCCAGTCACTGTGGAGTGCCTTATGCCCGCTCAGAGGCCCACTCAGAGAG 241
    |||||||
QY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
    |||||||
DB 242 CTGCTGGAGAGATATGTGACCGATGAGATGAGGAGACATGTGATCCTTCACC 301
    |||||||
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGluLeuAspLeu 100
    |||||||
DB 302 CATCGCAAGAACTACTACGTGTAGTGGCCGGAATGAGAAATCCAGTGAATGACCTA 361
    |||||||
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
    |||||||
DB 362 CAAGCATCCGATCGATGATATTACGCGACCCCTCAAGTTTCGTGTGAGAGCAT 421
    |||||||
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140
    |||||||
DB 422 GTGGAGGAATACAGAGATGAATCATTTCTTTCCGAGAGGCTGCAATGTTAA 481
    |||||||
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
    |||||||
DB 482 GACAAACTTTCAGTAAGCAGACAGATCTTGTGACCATGCCCTGCACATATCGCATGAT 541
    |||||||
QY 161 GluLeu 162
    |||||||
DB 542 GAGCTA 547
    |||||||

```

Search completed: March 31, 2003, 04:51:14  
Job time : 1041.65 secs



JOURNAL  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:954339

Seq primer: -40RP from Glbco  
High quality sequence stop: 459.  
Location/Qualifiers

## FEATURES

source

1. 549  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1548981"  
/clone\_id="Scars\_mammary\_gland\_NMLMG"  
/sex="female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified pT73 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 154 a 133 c 150 g 112 t

ORIGIN

## Alignment Scores:

Pred. No.: 6.28e-98 Length: 549  
Score: 832.00 Matches: 160  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 98.77% Mismatches: 0  
Query Match: 98.23% Indels: 0  
DB: 10 Gaps: 0

US-10-082-502-21 (1-162) x BE198865 (1-549)

OY 1 ArgatgserglnaspLeuHiscysgluAlaCysArgAlaLeuValaspGluLeuGluTrp 20  
DB 19 CGAAGGACCAAGATCTACACTGTGAGCTTGCAGAGGCTGTGTGATGAATTAAGTGG 78  
OY 21 GluileAlaArGValaspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 79 GAAATGCCCGCTGGACCCCAAGAGACCATTCAGATGGATCCTCCGAATCAATCCA 138  
OY 41 AspGlyserGlnserValaIgluValProTyrAlaArgserGluAlaHisleuThrGlu 60  
DB 139 GATGGCACCGATGATGTGTGAGGTACCTTATGCCCTCAGAAAGCCACCTCACAAG 198  
OY 61 LeuLeuGluGluValCysAspArgMetLysgluTyrGlyGluGlnIleaspProSerThr 80  
DB 199 TTGCTTGAGGAGGTGTGTGACGAAATGAAGAGTACGGGAAACAAATTAACCTTACC 258  
OY 81 HisArgLysAsnTyrValaArgValaIserArgAsnGlyLysSerGluLeuAspLeu 100  
DB 259 CACCGCAAGAACTAGTACGTACGCTGTGAGCCGGAATGCAGAAATCCAGTAAGACTTA 318  
OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 319 CAGGCAATCCGAATGATCATATATCAGCGGACCCCTCAAGTTGCGTGTAGAGCATT 378  
OY 121 ValGlnGlyTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 379 GTGGAAGAAATACGAGAGTACGACTTTCGAATCTTCTCCACAGAGGCTGACAACTTAA 438  
OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisasp 160  
DB 439 GACAAACTTTTCAGTAAACGAGACAGATCTATGTGACCAATGCCCTGCACAGATCTACGAT 498  
OY 161 GluLeu 162  
|||||

DB 499 GAGCTA 504

RESULT 15  
BM832828

LOCUS

DEFINITION

BM832828 588 bp mRNA linear EST 06-MAR-2002  
K-EST0107375 SSSNU484s1 Homo sapiens cDNA clone SSSNU484s1-11-B03  
5', mRNA sequence.

ACCESSION

BM832828

VERSION

BM832828.1 GI:19189237

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

TITLE

21C Frontier Korean EST Project 2001

JOURNAL

Unpublished (2002)

COMMENT

Contact: Kim YS

FEATURES

Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-353, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.krdb.re.kr  
Plate: 11 row: B column: 03  
High quality sequence stop: 588.  
Location/Qualifiers

source

1. 588

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="SSSNU484s1-11-B03"

/clone\_id="SSSNU484s1"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/cell\_line="SNU-484"

/lab\_host="Top10F"

/note="Organ: Stomach; Vector: pT718RP1; Site:1; EcoRI; Site:2; NOTI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and NOTI14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."

BASE COUNT 149 a 140 c 176 g 123 t

ORIGIN

## Alignment Scores:

Pred. No.: 1.26e-97 Length: 588  
Score: 830.00 Matches: 158  
Percent Similarity: 98.77% Conservative: 2  
Best Local Similarity: 97.53% Mismatches: 2

```

Db 258 CGAAGAGCCAGATCTACCTGTGAGGCTGTGTCGAGGCTGTGTCGATTAAGACTGC 317
QY 21 GtUilealArValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40
Db 318 GAAATTCGCCGCTGACCCCAAGAACATTCAGATGGGATCTTCCGATCAATCCA 377
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60
Db 378 GATGGCAGCAGTACGTTGGAGGTACCTTAATGCCCCGCTCAAGAGGCCACCTCACAGAG 437
QY 61 LeuLeuGlnGluValCysAspArgMetGlyGluTyrGlyLeuGlnIleAspProSerThr 80
Db 438 TTCTTGGAGAGGTGTGTGACCAATGAAGAGATACGGGAGAACAGATTGACCTTCTTACC 497
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGlnLeuAspLeu 100
Db 498 CACCGCAGAACTACGTAACCGCTGTGAGCGGAAATGGAATTCAGTAACTAAGACTTA 557
QY 101 GlnGlyTyrLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAla-CysGluSerIle 120
Db 558 CAGGCGATCCGAAATGATCAGATATCAGCGCACCCCTCAAGTTTGGCTGTGAGACAT 617
QY 120 eValGlnGluTyrGluAspGluLeuIleGlnPheSerArgGlnAlaAspAsnValIy 140
Db 618 TGTGGAAGATACGAGGATAGCTTATCGAATCTCTCCAGAGAGGCTACAAACGTTAA 677
QY 140 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAs 160
Db 678 AGCAAACTTGGCAGTAAGGGAGACAGATCTATGTGACCAATGCCCTGCACAGATCTCAGA 737
QY 160 pGluLeu 162
Db 738 TGAACCT 744

RESULT 13
BI411532 709 bp mRNA linear EST 14-AUG-2001
LOCUS BI411532
DEFINITION 602965126F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5120574.5',
mRNA sequence.
ACCESSION BI411532
VERSION BI411532.1 GI:15172455
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 709)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
JOURNAL NIH-MGC http://mgs.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldi, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
plate: LLAM11294 row: c column: 07
high quality sequence start: 4
high quality sequence stop: 701.
Location/Qualifiers
1. 709
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a

```

```

modified polylinker; Site.1: NotI; Site.2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTCACATCTGAACTGGAGAGCGGCGCTGTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldi.
BASE COUNT 191 a 170 c 214 g 132 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 6,76e-98 Length: 709
Score: 833.00 Matches: 159
Percent Similarity: 98.15% Conservative: 0
Best Local Similarity: 98.15% Mismatches: 3
Query Match: 98.35% Indels: 0
DB: 13 Gaps: 0
US-10-082-502-21 (1-162) x BI411532 (1-709)
QY 1 ArgArgSerGlnAspLeuHisCysGlyValCysArgAlaLeuValAspGluLeuGluTyr 20
Db 164 CGAAGAGCCAGATCTACCTGTGAGGCTGTGTCGAGGCTGTGTCGATTAAGACTGC 223
QY 21 GtUilealArValAspProLysThrIleGlnMetGlySerPheArgIleAspPro 40
Db 224 GAAATTCGCCGCTGACCCCAAGAACATTCAGATGGGATCTTCCGATCAATCCA 283
QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGlnAlaHisLeuThrGlu 60
Db 284 GATGGCAGCAGTACGTTGGAGGTACCTTAATGCCCCGCTCAAGAGGCCACCTCACAGAG 343
QY 61 LeuLeuGlnGluValCysAspArgMetGlyGluTyrGlyLeuGlnIleAspProSerThr 80
Db 344 TTCTTGGAGAGGTGTGTGACCAATGAAGAGATACGGGAGAACAGATTGACCTTCTTACC 403
QY 81 HisArgLysAsnTyrValArgValValSerArgAsnGlyLysSerGlnLeuAspLeu 100
Db 404 CACCGCAGAACTACGTAACCGCTGTGAGCGGAAATGGAATTCAGTAACTAAGACTTA 463
QY 101 GlnGlyTyrLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 464 CAGGCGATCCGAAATGATCAGATATCAGCGCACCCCTCAAGTTTGGCTGTGAGACAT 523
QY 121 ValGlnGluTyrGluAspGluLeuIleGlnPheSerArgGlnAlaAspAsnValIys 140
Db 524 GTGGAAGATACGAGGATAGCTTATCGAATCTCTCCAGAGAGGCTGCACAACTGTTAA 583
QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160
Db 584 GACAACTTGGCAGTAAGGGAGACAGATCTATGTGACCAATGCCCTGCACAGATCTCAGAT 643
QY 161 GluLeu 162
Db 644 GAGCTA 649

RESULT 14
BI198865 549 bp mRNA linear EST 26-JUN-2000
LOCUS BI198865
DEFINITION ug83d11.y1 Soares_mammary_gland_NLMJG Mus musculus cDNA clone
IMAGE:1548981.5' similar to WP.F01F1.11 CE01224 ;, mRNA sequence.
ACCESSION BE198865
VERSION BE198865.1 GI:8711034
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 549)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap/
COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

```

KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 678)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov  
Plate: LLM8523 row: e column: 01  
High quality sequence stop: 590.  
Location/Qualifiers  
1. 678  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_image="3486696"  
/clone\_lib="NCI-CGAP\_Mam5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1; Salt: Site:2; NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

BASE COUNT 192 a 154 c 191 g 141 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.69e-98 Length: 678  
Score: 834.00 Matches: 162  
Percent Similarity: 99.39% Conservative: 0  
Best Local Similarity: 99.39% Mismatches: 0  
Query Match: 98.47% Indels: 1  
DB: 10 Gaps: 0

US-10-082-502-21 (1-162) x BE309953 (1-678)

OY 1 ArgatgserglnaspleuHiscysglialacysargalaleuvalaspluleuglttrp 20  
DB 33 CGAGGAGCCAGATCTACACTGTGTGAGCTGCAGGGCTCTGTGATCAATTAGAGTGG 92  
OY 21 Gtlualealargvalaspprolaylsrthrlleglmetglyserphearglleaspro 40  
DB 93 GAAATGTGCCCGGTGAGCCCAAGAACATTCAGATGGGATCCTTCGATCAATCA 152  
OY 41 Aspglyserglnservalaigluvalprotyralaargserglualahisleuthrglu 60  
DB 153 GATGCACCCACATCACTGTTGTGAGGTACTTATGCCCTCAGAGGCCACCTCACAGAG 212

OY 61 LeuleuglgluvalaCysaspargmetlysglytrgylgluglnlilaaspproserrhr 80  
DB 213 TTGCTTGAGAGAGGTGTGTGACCGAATGAAGAGTACGGGGAACAGATTGACCTTCTAC 272  
OY 81 HlsatrglyasnytrvalarvalaValserargasnlyglusersercluleuaspleu 100  
DB 273 CACCGCAAGAACTAGTACGGCTGTGAGCCGGAATGGAGATCCAGTAACTAGACTTA 332  
OY 101 Glnglylleatrglleaspserrasplleserglythrleuvalpheaalcysglusertle 120  
DB 333 CAGGCGATCCGAATGACTCAGATATCAGCGGACCCCAAGTTGGCGTGTGAGAGCATTT 392  
OY 121 Valgluglutyrgluaspluleuilegluphepserargglualahaspasnavallys 140

DB 393 GTGGAAGATACGAGGATAGCTTATCGAATTCCTCCAGAGAGCTACACGTTAAA 452  
OY 141 Asplylsleucysesser-lyargthrAspleucysasphlisaaleuHlsarSerrHisas 160  
DB 453 GACAAACTTGTGAGTAAGCGGACAGATCTATGTACCACTGCCCTGCACAGATTCACGA 512  
OY 160 pgluleu 162  
DB 513 TGAGCTA 519  
RESULT 12  
BI408381  
LOCUS  
DEFINITION 602964013p1 NCI-CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5119563 5', mRNA sequence.  
ACCESSION BI408381 GI:15169304  
VERSION BI408381.1 GI:15169304  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 891)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov  
Plate: LLM11291 row: 1 column: 04  
High quality sequence stop: 873.  
Location/Qualifiers  
1. 891  
/organism="Mus musculus"  
/strain="CZECH IT"  
/db\_xref="taxon:10090"  
/clone\_image="5119563"  
/clone\_lib="NCI-CGAP\_Lu33"  
/tissue\_type="pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pTRT3D-Pac (Pharmacia) with a modified polylinker; Site:1; NotI; Site:2; EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5', TGTATCAATCTGAAGTGGAGCGCGCGCTGTGTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTRT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 240 a 215 c 258 g 178 t  
ORIGIN

Alignment Scores:  
Pred. No.: 7e-98 Length: 891  
Score: 834.00 Matches: 162  
Percent Similarity: 99.39% Conservative: 0  
Best Local Similarity: 99.39% Mismatches: 0  
Query Match: 98.47% Indels: 1  
DB: 13 Gaps: 0

US-10-082-502-21 (1-162) x BI408381 (1-891)

OY 1 ArgatgserglnaspleuHiscysglialacysargalaleuvalaspluleuglttrp 20  
DB 33 CGAGGAGCCAGATCTACACTGTGTGAGCTGCAGGGCTCTGTGATCAATTAGAGTGG 92

Pred. No.: 3,83e-98 Length: 724  
 Score: 835.00 Matches: 159  
 Percent Similarity: 98.77% Conservative: 1  
 Best Local Similarity: 98.15% Mismatches: 2  
 Query Match: 98.58% Indels: 0  
 DB: 14 Gaps: 0

US-10-082-502-21 (1-162) x BQ210416 (1-724)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 629 CGAAGAGAGCAGATCTACACTGTGGAGGCTGTGGTGAATTAAGAGTGG 570.  
 QY 21 GluIleAlaArgValAspProLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 |||||  
 DB 569 GAAATTCCTCCGCTGACCCCAAGAACCATTCAGATGGATCTCCGAATCAATCCA 510  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 509 GATGGCAGCCAGTCAGTGTGGAGGTCCTTATGCCCCGCTCAAGAGCCCTCACGGAG 450  
 QY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||  
 DB 449 TTGCTTGAGAGATATGTGACCAATGAAGAGATATGGGAACAGATGACCTTCTACC 390  
 QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 100  
 |||||  
 DB 389 CACCGCAAGAACTACGTCCTGTTGTCGGCCGAGACGAGATCCAGTCAATCAAGACTTA 330  
 QY 101 GlnGlyTyrLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 329 CAGGGATATCCGAATTCAGATGATATCAGAGGACCTCAAGTTGGTGTGAGAGCATT 270  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 269 GTGAGAGAGTACAGAGATGAACTATTGATTTCTTTCAAGAGAGCGTGAACGTTAA 210  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 DB 209 GACAACTTTGACAGTAAAGCGACAGATCTATGTACCATCTCCCTGCACAGATCTCATGAT 150  
 QY 161 GluLeu 162  
 |||||  
 DB 149 GAGCTA 144

RESULT 10 802 bp mRNA linear EST 01-APR-2001  
 LOCUS BF608664  
 DEFINITION MYL\_00106 Mouse 9-day fetus cDNA library ICRFP522 Mus musculus  
 accession BF608664  
 version BF608664.1 GI:13505156  
 keywords EST.  
 source house mouse.  
 organism Mus musculus.  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 802)  
 Yahyawi, M., Hennig, S., Neidhardt, L., Radloff, U., Hermann, B.G.,  
 Lehnach, H. and O'Brien, J.  
 Detection of a high number of novel genes in a 9-day mouse embryo  
 cDNA library normalised by oligonucleotide fingerprinting  
 unpublished (2001)  
 JOURNAL COMMENT  
 Contact: Hennig S  
 Laboratory 123, dept. Lehnach  
 Max-Planck-Institut fuer Molekulare Genetik  
 Ihnestr. 63-73, D-14195 Berlin, Germany  
 Tel: +49 30 8413 1612  
 Fax: +49 30 8413 1380  
 Email: hennig@molgen.mpg.de  
 EST's are made from clones being representatives of clone clusters.  
 Clone clusters were calculated from oligonucleotide fingerprints.  
 PCR Primers

FEATURES  
 source  
 1.802  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="ICRF522B2440"  
 /clone.lib="Mouse 9-day fetus cDNA library ICRFP522"  
 /tissue.type="whole embryo"  
 /dev\_stage="embryonic 9-day"  
 /lab\_host="E.coli, XL1 blue"  
 /note="Vector: pSVSPORT1; Site\_1: NotI; Site\_2: SalI;  
 library preparation by oligo-dT priming of RNA. Clones can  
 be ordered from the Resource Center in Berlin,  
 http://www.rzpd.de."

BASE COUNT 193 a 206 c 243 g 159 t 1 others  
 ORIGIN

#### Alignment Scores:

Pred. No.: 4.45e-98 Length: 802  
 Score: 835.00 Matches: 160  
 Percent Similarity: 99.38% Conservative: 1  
 Best Local Similarity: 98.77% Mismatches: 1  
 Query Match: 98.58% Indels: 0  
 DB: 12 Gaps: 0

US-10-082-502-21 (1-162) x BF608664 (1-802)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
 |||||  
 DB 300 CGAAGAGAGCAGATCTACACTGTGGAGGCTTGCAGGCTGTGGTGAATTAAGAGTGG 359  
 QY 21 GluIleAlaArgValAspProLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
 |||||  
 DB 360 GAAATTCCTCCGCTGACCCCAAGAACCATTCAGATGGATCTCCGAATCAATCCA 419  
 QY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
 |||||  
 DB 420 GATGGCAGCCAGTCTGTGGAGTACCTTATGCCCCGCTCAAGAGCCCTCACAGAG 479  
 QY 61 LeuLeuGluGluValAlaCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 |||||  
 DB 480 TTGCTTGAGAGAGTGTGTGACCAATGAAGAGTACGTTGACAGATTCACCTCTACC 539  
 QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerSerGluLeuAspLeu 100  
 |||||  
 DB 540 CACCGCAAGAACTACAGTACGCTGTGAGCCGGAATGGAATCCAGTGAATCAACTTA 599  
 QY 101 GlnGlyTyrLeuArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
 |||||  
 DB 600 CAGGGCATCCGAATTCAGATGATATCAGCGGACACCTCAAGTTTGGTGTGAGAGCATT 659  
 QY 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140  
 |||||  
 DB 660 GTGGAAGAAATACAGAGATGAGCTTATGCAATCTCTCCAGAGAGCGTGANCACTTAA 719  
 QY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 |||||  
 DB 720 GACAACTTTGACAGTAAAGCGACAGATCTATGTGACCATGCTCCGTCAGACATCTCAGAT 779  
 QY 161 GluLeu 162  
 |||||  
 DB 780 GAACCT 785

RESULT 11 678 bp mRNA linear EST 26-OCT-2000  
 LOCUS BE309953  
 DEFINITION 601091903F1 NCI\_CGAP\_Mam5 Mus musculus cDNA clone IMAGE:3486696 5',  
 accession BE309953  
 version BE309953.1 GI:9168932

and the oligo-dT track served to verify it as a clone from the  
normalized duodenum library cDNA Library Preparation: M.B. Soares  
Lab Clone distribution: clones will be available through Research  
Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA-Yes.

# FEATURES

Source location/Qualifiers  
1..674  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-EB1-clone-b-02-0-UI"  
/clone\_lib="UI-R-EB1"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI; UI-R-EB1 is a  
non-normalized Rat Duodenum Library (RDL) constructed in  
pT737 PAC vector according to the procedure described by  
Bonaldi, Lennon & Soares (Genome Research 6:  
791-806, 1996). The oligonucleotide used to prime first  
strand synthesis contained the sequence tag TGTGGTCAT  
between the Not I cloning site and dT18 stretch. The Rat  
Duodenum tissue was provided by Tom Freeman of the Sanger  
Center.  
TAG\_LIB=UI-R-EB1  
TAG\_TISSUE=duodenum  
TAG\_SEQ=TGTGGTCAT"

## Alignment Scores:

Pred. NO.: 3.45e-98 Length: 674  
Score: 835.00 Matches: 159  
Percent Similarity: 98.77% Conservative: 1  
Best Local Similarity: 98.15% Mismatches: 2  
Query Match: 98.58% Indels: 0  
DB: 14 Gaps: 0

US-10-082-502-21 (1-162) x BQ202181 (1-674)

0Y 1 ArgATGserGlnAspLeuHIScysAGTValAcysArgAlaLeuValAspGluLeuGluTrp 20  
DB 636 CGAAGAGACCCAGAGATCTCACTGTGAGCTTGCAAGGCTGTGGATGATTAAGTGG 577  
0Y 21 GluIleAlaArgValAspProLysThrIleGlnMetGlySerPheArgIleAsnPro 40  
DB 576 GAAATGTCCCGGTGGAGCCCAAGAGACCAATTCAGATGGATCTCTCCGAATCAATCCA 517  
0Y 41 AspGlySerGlnSerValAlaGluValProGlyAlaArgSerGluAlaHisLeuThrGlu 60  
DB 516 GATGGCAGCCAGTCAGTTGTGTGAGGTTCTTATGCCCTCAGAGGCCACCTCAGGAG 457  
0Y 61 LeuIleGluGluValCysAspArgMetLysGluGluGluGluGluGluGluGluGluGlu 80  
DB 456 TTGCTTGGAGATATGTGACCGAATGAAGATGTGGGAACACATTTGACCTTTCTAC 397  
0Y 81 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 100  
DB 396 CACGCAAGAACTACGTCGCTGTGTGCGCGGAAGGAGATCCAGTAAGTAAGTAA 337  
0Y 101 GlnGlyIleArgGlyLeuAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 336 CAGGATATCCCAATTCAGATTCAGATTCAGTGCACCCCAAGTTGCTGTGAGAGCAT 277  
0Y 121 ValGluGluTyrGluAspGluLeuLeuPhePheSerArgGluAlaAspAsnValLys 140  
DB 276 GTGGAGGATACGAGAGATGAATCTATTGAATCTTTTAAAGAGAGCTGACAAACGTTAA 217  
0Y 141 AspLysLeuCysSerLysArgThrAspLeuLysAspHisAlaLeuHisArgSerHisAsp 160  
DB 216 GACAAACTTTGCAAGTACGAGGACGAGATCTATGTGACATGCCCTGCACAGATCTCATAT 157

0Y 161 GluLeu 162  
DB 156 GAGCTA 151

## RESULT 9

LOCUS BQ210416/c 724 bp mRNA linear EST 02-MAY-2002

DEFINITION UI-R-DY1-col-e-06-0-UI-s1 UI-R-DY1 Rattus norvegicus cDNA clone

ACCESSION BQ210416  
VERSION BQ210416.1 GI:20426881

KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 724)

AUTHORS Bonaldi,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to verify it as a clone from the  
normalized cartilage library cDNA Library Preparation: M.B. Soares  
Lab Clone distribution: clones will be available through Research  
Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA-Yes.

## FEATURES

Source location/Qualifiers  
1..724  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-DY1-col-e-06-0-UI"  
/clone\_lib="UI-R-DY1"  
/issue\_type="Cartilage"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Femur and Tibia; Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker. Site 1: Not I;  
Site 2: Eco RI; UI-R-DY1 is a normalized cDNA library  
containing the following tissue(s): Rat Cartilage from  
Femur and Tibia. The library was constructed according to  
Bonaldi, Lennon & Soares, Genome Research, 6:791-806,  
1996. First strand cDNA synthesis was primed with an  
oligo-dT primer containing a Not I site. Double stranded  
cDNA was ligated to an EcoR I adaptor, digested with Not  
I, and cloned directionally into pT73D-Pac vector. The  
oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is CTATGGAGC. The Rat  
cartilage tissue was provided by Dr Jeff Stevens at the  
University of Iowa.  
TAG\_LIB=UI-R-DY1  
TAG\_TISSUE=cartilage  
TAG\_SEQ=CTATGGAGC"

BASE COUNT 156 a 206 c 157 g 205 t

## ORIGIN

Alignment Scores:

|||||  
Db 395 GATGGACAGCCAGTTCAGTGTGGAGGATACCTTATGCCCGCCACAGAGCCACCTCACAGAG 454  
OY 61 LeuLeuGluGluValGlyAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
Db 455 TTCTGTGAGGAGGTGTGTGACCAATGAAGATACGGGGAACAGATGACCTTCTAC 514  
OY 81 HisArgLysAsnTyrValArgValIleSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
Db 515 CACCGCAAGAACTACGATACCGCTCGTAGCGGAAATGGAAATCCAGTAGACACTTA 574  
OY 101 GlnGlyTleArgTleAspSerAspIleSerGlyTyrLeuLysPheAlaCysGluSerIle 120  
Db 575 CAGGCACTCCGAATTCAGATTCAGATTCAGCGCACCTCAAGTTGCGGTGAGAGCAT 634  
OY 121 ValGlnGluTyrGluAspGluLeuIleGluPhePheSerArgGluIleAspAsnValLys 140  
Db 635 GTGGAAAGAAATACAGAGATGAGTTCATCGAATTCCTTCACAGAGAGGCTGACAAAGTTAAA 694  
OY 141 AspLysLeuCysSerLysArgTyrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
Db 695 GACAACTTTCAGTAGACGACAGATCATGTGACCATGCCCTGCACAGATCTCAGCAT 754  
OY 161 GlnLeu 162  
Db 755 GAGCTA 760  
RESULT 7  
LOCUS AW919569 665 bp mRNA linear EST 25-MAY-2000  
DEFINITION Rattus norvegicus cDNA clone RGIzf11 5' end, mRNA sequence.  
ACCESSION AW919569  
VERSION AW919569.1 GI:8085366  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 665)  
AUTHORS Lee,N.H., Gloddek,A., Chandra,I., Mason,T.M., Quackenbush,J.,  
Kerlavage,A.R. and Adams,M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat  
Gene Index  
JOURNAL Unpublished (1998)  
COMMENT Contact: Lee, NH  
The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
This clone is available through the ATCC, contact the ATCC  
tel#03-365-2700 for further information  
Seq primer: M13 Reverse.  
FEATURES  
source  
1. 665  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone\_lib="RGIzf11"  
/clone\_id="Rat gene index, normalized rat, norvegicus,  
Bento Soares"  
/tissue\_type="mix - brain, ovary, placenta, kidney, lung,  
liver, embryo, heart, muscle, spleen"  
/lab\_host="SOLR"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; Estimated insert size approx.1 Kb"

Percent Similarity: 98.77%  
Best Local Similarity: 98.15%  
Query Match: 98.58%  
DB: 10 Gaps: 0  
US-10-082-502-21 (1-162) x AW919569 (1-665)  
OY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTyr 20  
Db 620 CGAAGAGACCGAGATCTACTGAGCTTGCAGGGCTCTGTGATGATTAAGAGTGG 561  
OY 21 GlnIleAlaArgValAspProLysLysThrIleGlnMetCysSerPheArgIleAsnPro 40  
Db 560 GAATATGCCCGCTGAGACCCCAAGAGACCATTCAGATGGGATCTTCCAAATCAATCCA 501  
OY 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluLanHisLeuThrGlu 60  
Db 500 GATGGACAGCCAGTCAGTGTGGAGGTTCTTATGCGCCGCTCAGAGGCCACCTCACGGAG 441  
OY 61 LeuLeuGluGluValGlyAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
Db 440 TTCTGTGAGGAGATATGTGACGAAATGAAGATATGGGAAACAGATGACCTTCTAC 381  
OY 81 HisArgLysAsnTyrValArgValIleSerArgAsnGlyGluSerSerGluLeuAspLeu 100  
Db 380 CACCGCAAGAACTACGCTCCTGTGTCGCCGACGAGAGATCCAGTAACATGACTTA 321  
OY 101 GlnGlyTleArgTleAspSerAspIleSerGlyTyrLeuLysPheAlaCysGluSerIle 120  
Db 320 CAGGATATCCGAATTCGATTCGATTCAGTTCAGCACCTCAAGTTTCTGTGAGACAT 261  
OY 121 ValGlnGluTyrGluAspGluLeuIleGluPhePheSerArgGluIleAspAsnValLys 140  
Db 260 GTGGAGAGATACAGAGATGAATTAATGATTTCTTTCAAGAGAGCTGACAAAGTTAAA 201  
OY 141 AspLysLeuCysSerLysArgTyrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
Db 200 GACAACTTTCAGTAGACGACAGATCATGTGACCATGCCCTGCACAGATCTCAGCAT 141  
OY 161 GlnLeu 162  
Db 140 GAGCTA 135  
RESULT 8  
LOCUS B0202181 674 bp mRNA linear EST 02-MAY-2002  
DEFINITION UT-R-EB1-cle-b-02-0-UI-61 UT-R-EB1 Rattus norvegicus cDNA clone  
UT-R-EB1-cle-b-02-0-UI 3', mRNA sequence.  
ACCESSION B0202181  
VERSION B0202181.1 GI:20418646  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 674)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msosares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site



ORGNISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
10349636

2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiyagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Kataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleschman, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schiraldi, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustlich, S., Hill, D., Hofmann, M., Hume, D. A., Kamliya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

5 (bases 1 to 1255)  
Aaach, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Aizawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imocani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Oza, C., Quackenbush, J., Schiraldi, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (18-AUG-2000) Yoshinide Hayashizaki, The Institute of

## COMMENT

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGATTCGACGTTAAATTAATTAATTCACCCCCCCC 3']. cDNA was prepared by using triethanolamine-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot - 3.0 and subtraction to Rot - 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATTCGACGTTAAATTAATTAATTCACCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from lambda Phi I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

## FEATURES

## source

Location/Qualifiers

1..1255  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM:DB:5330432A10"  
/db\_xref="MGD:MGI:1899093"  
/db\_xref="taxon:10090"  
/clone="5330432A10"  
/sex="male"  
/tissue\_type="pituitary gland"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_strage="adult"  
1..1255  
/gene="Tmem4"  
215..763  
/gene="Tmem4"  
/note="data source:MGD, source key:MGI:1928477, evidence:ISS  
putative  
transmembrane protein 4"  
/codon\_start=1  
/protein\_id="BAB31921.1"  
/db\_xref="GI:12860333"  
/db\_xref="MGD:MGI:1928477"  
/translation="MKKQWALALLCVLTGTANARKSDIHCGRALVDELEIAR VDKRTIOWGSEFRINPDGSSVVEVPAREAHITTELEVCDDMKKEGQIDPSTHR KNVYRVVRSNGESELDDGIRIDSDISGTLKFACEISIVEYEDLETFESREADNVK DKLSKRTDIDCDHRLHRSDEL"

## CDS

## gene

BASE COUNT 309 a 293 c 354 g 299 t  
ORIGIN

## Alignment Scores:

Pred. NO.: 2.35e-99 Length: 1255  
Score: 847.00 Matches: 162  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-082-502-21 (1-162) x AK019927 (1-1255)

OY 1 Argtrgserglnaaprylenhiscygluylacysargalalevalaarglileuglutr 20  
DB 275 CGAAGGACCAAGATCTACACTGTGACCTTCGAGCGCTCGGAGATTAAGAGCG 334  
OY 21 gtuilealaargvalaaprolusylthrilleglnmetgllyserphargiilaappro 40  
DB 335 GAAATTCGCCCGCGGAGCCCAAGAAAGCACTTCAGATGGATCTCCGAAATCAATCA 394  
OY 41 Aspelyserglnservalalgluvalprotyrgalargserglualanlsleutthclu 60





OY 161 Glut1eu 162  
 DB 659 GACCTT 664

RESULT 4  
 BI412539

LOCUS BI412539 843 bp mRNA linear EST 14-AUG-2001  
 DEFINITION 60299095F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5146967 5',  
 mRNA sequence.

ACCESSION BI412539  
 VERSION BI412539.1 GI:15173462  
 KEYWORDS EST.

SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 843)  
 TITLE NIH-MGC http://mgc.nci.nih.gov/  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LAM1362 row: n column: 24  
 High quality sequence start: 38  
 Location/Qualifiers  
 1. 843  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5146967"  
 /clone\_1id="NCI\_CGAP\_Lu33"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: lung; Vector: pTZ19-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5',  
 TGTTCACATCTGAGAGGAGCGCGCTGTCTTTTCTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pTZ19 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 239 a 195 c 238 g 171 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,31e-99 Length: 843  
 Score: 847.00 Matches: 162  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-082-502-21 (1-162) x BI412539 (1-843)

OY 1 ArgArgSerGlnAspLeuHIScysGlyAlaCysArgAlaLeuValAspGlnLeuGlnTyr 20  
 DB 214 CGAAGGAGCCAAAGATCTACACTGTGAGCTTGCAGGCGCTCGGATGTAATTGAGCTGC 273

OY 21 GluIleAlaArgValAspProLysLysThrIleGlnMetGlySerPheArgIleAspPro 40  
 DB 274 GAAATTGCCCGGTGGAGCCCAAGAGACCAATTCAGATGGATCTTCCGAAATCAATCA 333

OY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaIleLeuThrGlu 60  
 DB 334 GATGGACAGCCAGTCAGTGTGGAGGTACTTATGGCCCTCAGAGGCCACCTCAGACAG 393

OY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80  
 DB 394 TTGCTGTAGGAGGTGTGTGACCCGAATGAAAGAGGAGGGAACAGATTGACCTTCTACC 453

OY 81 HisArgLysAsnTyrValArgValAlaSerArgAsnGlyGlnSerSerGluLeuAspLeu 100  
 DB 454 CACCGCAAGAACTACGTACGGCTGTGACCGCGAATGAGAAATCAGGAACCTAGACTTA 513

OY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGlnSerIle 120  
 DB 514 CAGGCAATCCGAATTTGACTCAGATATCAGCGCACCCCTCAAGTTTGGCTGTGAGAGCAT 573

OY 121 ValGluGluTyrGluAspGlnLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
 DB 574 GTGGAAAGAAATCGAGGATGACCTTATCGAAATCTTCTCCAGAGAGGCTGACACAGTTTAA 633

OY 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
 DB 634 GACAACTTTCAGATGACGAGACAGATATATGACCTGCCCTGCACAGATTCACGAT 693

OY 161 Glut1eu 162  
 DB 694 GACCTA 699

RESULT 5  
 AK013014

LOCUS AK013014 853 bp mRNA linear HTC 19-JAN-2002  
 DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length  
 enriched library, clone:2810406124:transmembrane protein 4, full  
 insert sequence.

ACCESSION AK013014  
 VERSION AK013014.1 GI:12850114  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain: C57BL/6J) 10, 11 days embryo cDNA to mRNA,  
 clone\_1id:RIKEN full-length enriched mouse cDNA library  
 clone:2810406124.

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
 Carninci, P., and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE 99279253  
 PUBMED 10349636

REFERENCE 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE 20499374  
 PUBMED 11042159

REFERENCE 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishi, K., Kitsuhei, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multichannel sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE 20530913  
 PUBMED 11076861

REFERENCE 4  
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

/db\_xref="MCD:MGI:1928477"  
/translation="MKGMWLLALLGVLLGTAMARRSODLHGCACRALVDELEWEIAR  
VPRKTIQMSFRINDPSQSYVEVPYARSEAHTELEEVCDRMRKEGEQIDPSTHR  
KRYVRVRSNGSESLDLOGIRIDSDISGLTKFACSEIVEEDELIEFFSEADNVK  
DKLGRMDLCOHALHRSDEL"

BASE COUNT 210 a 180 c 225 g 157 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.15e-99 Length: 772  
Score: 847.00 Matches: 162  
Percent Similarity: 100.00% Conserved: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-082-502-21 (1-162) x AK00914 (1-772)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 164 CGAAGAGCCGCAAGATCTACACTGTGAGCTTCGAGGCGCTCGTGATTAAGAGTGG 223  
QY 21 GlnAlaAlaArgValAspProLysTrpThrIleGlnMetCysSerPheArgIleAsnPro 40  
DB 224 GAATATGCCCGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGAATCAATCCA 283  
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 284 GATGGCAGCCAGTGTGGAGGATACCTTATGCCCGTCAGAGGCCACCTCACAGAG 343  
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80  
DB 344 TTCTCTGAGAGGAGTGTGTGACCAATGAAGAGTACGCGGACAGATTCACCTTCAC 403  
QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 100  
DB 404 CACCGCAAGACTACGATCCGTCGTGAGCCGGAATGAGAAATCCAGTACTAGACTTA 463  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 464 CAGGGCATCCGAATTCAGATTCAGATTCAGGAGACCTCAAGTTGGCTGTGAGAGCAT 523  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 524 GTGGAAGAAATACAGAGATGAGCTTATTCGAATCTTCTCCAGAGAGGCTGCAACGTTAAA 583  
QY 141 AspLysLeuLysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
DB 584 GACAACTTTCAGATTAAGCGGACAGATCTATGTGACCATCCCTGACACAGATCTCAGAT 643  
QY 161 GluLeu 162  
DB 644 GAGCTA 649

RESULT 3  
BI408525  
LOCUS 602964995F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5120471 5',  
DEFINITION mRNA sequence.  
ACCESSION BI408525  
VERSION BI408525.1 GI:15169448  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 808)  
NIR-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgarbs-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1293 row: n column: 24  
High quality sequence start: 4  
High quality sequence stop: 768.  
Location/Qualifiers

FEATURES  
source

1.808  
/organism="Mus musculus"  
/strain="CZECH II"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5120471"  
/clone\_1lb="NCI\_CGAP\_Lu33"  
/tissue\_type="Pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pTZ19-Pac (Pharmacia) with a  
modified polylinker. Site 1: NotI; Site 2: EcoRI; 1st  
strand cDNA was prepared from mRNA obtained from pooled  
lung tumors with a Not I - oligo(dT) primer [5'  
TCTTACCACTGTGAAGTGGAGAGCGCGGCTCTGTTTCTTTTCTTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified pTZ19 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 224 a 181 c 237 g 166 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.23e-99 Length: 808  
Score: 847.00 Matches: 162  
Percent Similarity: 100.00% Conserved: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-082-502-21 (1-162) x BI408525 (1-808)

QY 1 ArgArgSerGlnAspLeuHisCysGlyAlaCysArgAlaLeuValAspGluLeuGluTrp 20  
DB 179 CGAAGAGCCGCAAGATCTACACTGTGAGCTTCGAGGCGCTCGTGATTAAGAGTGG 228  
QY 21 GlnAlaAlaArgValAspProLysTrpThrIleGlnMetCysSerPheArgIleAsnPro 40  
DB 239 GAATATGCCCGGTGGACCCCAAGAACCATTCAGATGGGATCTTCCGAATCAATCCA 298  
QY 41 AspGlySerGlnSerValAlaGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60  
DB 299 GATGGCAGCCAGTGTGGAGGATACCTTATGCCCGTCAGAGGCCACCTCACAGAG 358  
QY 61 LeuLeuGluGluValCysAspArgMetLysGluTyrGlyGlnIleAsnProSerThr 80  
DB 359 TTCTCTGAGAGGAGTGTGTGACCAATGAAGAGTACGCGGACAGATTCACCTTCAC 418  
QY 81 HisArgLysAsnTyrValAlaArgValAlaSerArgAsnGlyLysSerGluLeuAspLeu 100  
DB 419 CACCGCAAGACTACGATCCGTCGTGAGCCGGAATGAGAAATCCAGTAACTAGACTTA 478  
QY 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120  
DB 479 CAGGGCATCCGAATTCAGATTCAGATTCAGGAGACCTCAAGTTGGCTGTGAGAGCAT 538  
QY 121 ValGluGluTyrGluAspGluLeuIleGluPheSerArgGluAlaAspAsnValLys 140  
DB 539 GTGGAAGAAATACAGAGATGAGCTTATTCGAATCTTCTCCAGAGAGGCTGCAACGTTAAA 598  
QY 141 AspLysLeuLysSerLysArgThrAspLeuCysAspHisAlaLeuHisArgSerHisAsp 160  
DB 599 GACAACTTTCAGATTAAGCGGACAGATCTATGTGACCATCCCTGACACAGATCTCAGAT 658

Db	385	CACCCCAAGAACACTAGCTACGCGCTCCTGTAGGCCGGAATGGAGATCCAGTAACTAAGACTTA	444
Oy	101	GNGJYLIEATRIIEASPSERAPLIESERGTYTHRLIULYSPHALCYSGLUSERTLE	120
Db	445	CAGGCATCCGAAATTGACCTACGATATCACAGCCACCCTCAAGTTTGCGGTAGAGCATTT	504
Oy	121	VALIGUGUTYRGLUASPGLUDEUILIEGLUPHESPESERATGUALAASPANVALYS	140
Db	505	GTGGAGAATAACGAGAGATGACCTTATCGAATTCCTTCACAGAGAGCGTCACAACGTTAAA	564
Oy	141	AsplysielucCysserilyarqphrnspleucysasaphislaaleuhilargseritasp	160
Db	565	GACCAACTTGTGATGAAGGAGACATATCTATGTGACACATGCCCTCACAGATCTGACGAT	624
Oy	161	Glueu 162	
Db	625	GAGCTA 630	
RESULT 2			
LOCUS	AK007914	772 bp	mRNA linear HTC 19-JAN-2002
DEFINITION		Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810060D19;transmembrane protein 4, full insert sequence.	
ACCESSION	AK007914		
VERSION	AK007914.1	GI:12841773	
KEYWORDS	HTC; CAP trapper.		
SOURCE		Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to mRNA, clone.lib.RIKEN full-length enriched mouse cDNA library clone:1810060D19.	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
AUTHORS		1 Carninci,P. and Hayashizaki,Y.	
JOURNAL MEDLINE PUBLISHED		High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)	
PUBLISHED		99279253	
REFERENCE		10349636	
AUTHORS		2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL MEDLINE PUBLISHED		Genome Res. 10 (10), 1617-1630 (2000)	
PUBLISHED		20499374	
REFERENCE		11042159	
AUTHORS		3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishik,K., Kitunai,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasihagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Watabiki,M., Yokeda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuzoe,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.	
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexillary sequencer	
JOURNAL MEDLINE PUBLISHED		Genome Res. 10 (11), 1757-1771 (2000)	
PUBLISHED		20530913	
REFERENCE		11076861	
AUTHORS		4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Kono,H., Adachi,T., Fukuda,S., Atkawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Salto,R., Kadote,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochwa,H., Knehl,P., Lewis,S., Matsuo,Y., Mikado,I., Pesole,G., Quackenbush,J., Schmitt,L.M., Steinhil,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boftell,D., Bojunga,N., Carinelli,P., de Bonaldo,M.F., Brownstein,M.J., Bull,C.,	

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 . . . .  
 Fletcher,C., Fujita,A., Gariboldi,M., Gustincich,S., Hill,D.,  
 Hotemann,M., Hume,D.A., Kamyay,M., Lee,N.H., Lyons,P.,  
 Marchionni,L., Mashima,T., Mazzarelli,J., Mombaerts,P., Notdone,P.,  
 Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
 Satou,K., Schonbach,C., Seya,T., Shibata,Y., Scorch,K.F., Suzuki,H.,  
 Toyono-Oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,  
 Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohstall,S.  
 and Hayashizaki,Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 21085660  
 11217851  
 5 (bases 1 to 772)  
 Aachak,J., Alzava,R., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
 Atakewa,T., Baldarelli,R., Bono,H., Brownstein,M., Butl,C.,  
 Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T.,  
 Haru,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F.,  
 Hume,D., Imotani,K., Ishii,Y., Itoh,H., Izawa,M., Kesukawa,T.,  
 Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,  
 Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,  
 Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,  
 Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,  
 Schirml,L., Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T.,  
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
 Tanaka,I., Tejima,T., Toya,T., Yamamura,T., Yamanaka,I.,  
 Yananishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and  
 Hayashizaki,Y.  
 Direct Submission  
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp/  
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
 Fax:81-45-503-9216)  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
 further details.  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues. First strand cDNA was primed with a primer  
 [5' GAGAGAAGAGCGCCGCAATTAATCTCAGTATTAAATAATACCCCCCCC 3']. cDNA was  
 prepared by using trehalose thermo-activated reverse transcriptase  
 and subsequently enriched for full-length by cap-trapper. cDNA went  
 through one round of subtraction to Rot - 20.0. Second strand cDNA  
 was prepared with the primer adapter of sequence[5'  
 GAGAGAAGAGCGCCGCAATTAATCTCAGTATTAAATAATACCCCCCCC 3']. cDNA  
 was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3'  
 end: SstI. Host: SOLR.  
 FEATURES  
 Source  
 location/Qualifiers  
 1..772  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM\_DB:1810060D19"  
 /db\_xref="MGD:MGI:1897650"  
 /db\_xref="taxon:10090"  
 /clone="1810060D19"  
 /sex="male"  
 /tissue\_type="pancreas"  
 /clone\_id="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="10 day old"  
 1..772  
 /gene="Tmem4"  
 104..652  
 /gene="Tmem4"  
 /note="data source:MGD, source key:MGI:1928477,  
 evidence:ISS  
 putative  
 transmembrane protein 4"  
 /codon\_start=1  
 /protein\_id="BAB25346.1"  
 /db\_xref="GI:12841774"

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL  
MEDLINE  
PUBMED  
99279253  
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL  
MEDLINE  
PUBMED  
20493974  
11042159

3 Shibata, K., Itoh, M., Alizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaiyagi, K.,  
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multiplexed sequencer  
Genome Res. 10 (11), 1757-1774 (2000)

JOURNAL  
MEDLINE  
PUBMED  
20530913  
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arkawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Alizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Salto, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Salto, R.,  
Kaoto, K., Matsuda, H., Ashburner, M., Batalov, S., Casavola, T.,  
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,  
Knehl, P., Lewis, S., Matsuo, Y., Stabli, F., Suzuki, R., Tomita, M.,  
Okazaki, Y., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustinich, S., Hill, D.,  
Hotmann, M., Hume, D. A., Kamlay, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombert, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyo-Oka, K., Wang, K. H., Weitz, C., Whiteaker, C., Wilming, L.,  
Wysshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kontsuki, S.  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851

5 (bases 1 to 754)  
Adachi, J., Alizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Arkawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T.,  
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F.,  
Hume, D., Imoto, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
Kurihara, C., Matsuyama, T., Miyazaki, Y., Nishi, K., Nomura, K.,  
Numasaki, R., Ono, M., Okazaki, Y., Okido, T., Ota, C., Okazaki, Y.,  
Salto, H., Salto, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, Y.,  
Tanaka, T., Tejima, Y., Toya, T., Yamanura, T., Yamanaka, I.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,  
URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222,

	COMMENT
	Fax:81-45-503-9216) Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details.
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGACGACGATCCAGACCTCTTTTCTTTTTTCTTAA 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GGAGAGCAATTCTCGAGTTAATTAATAATGCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.
FEATURES	location/Qualifiers
source	1..754 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="PANTOM.DB:2900022B12" /db_xref="MGD:MGI:1896245" /db_xref="taxon:10090" /clone="2900022B12" /sex="male" /tissue_type="hippocampus" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult"
gene	1..754 /gene="Ymen4" .633 /note="data source:MGD, source key:MGI:1928477, evidence:ISS putative transmembrane protein 4" /codon_start=1 /protein_id="BAB28909.1" /db_xref="GI:12850980" /db_xref="MGD:MGI:1928477" /translation="MKMGWLTALLGLVLTGMARRSQDLHCACRALVDLEWEIAR VDPKATIOMGSFRINPDGSGSVVEVPARSBAHLTELEVCDBMKRYGEIDISTSRH KNRYRVSRNSESSELDLGRIDSODISGLTKFACESIVEYEDELIEFSREADNYK DKLSKRTRDLCDAHLNRSHDEL"
CDS	BASE COUNT 200 A 175 C 220 G 159 T ORIGIN
	Alignment Scores: Pred. No.: 1,11e-99 Length: 754 Score: 847.00 Matches: 162 Percent Similarity: 100.00% Conservative: 0 Best local Similarity: 100.00% Mismatchches: 0 Query Match: 100.00% Indels: 0 DB: 11 Gaps: 0
US-10-082-502-21 (1-162) x AK013568 (1-754)	
OY	1 ATGATGSeRGinAspLeuNHIScysGlYlaAcysArGaLaLeuValAspGluLeuGluTrp 20
Db	145 CGAAGGAGCGCAAGAATCTACACTGTGGAGCTTGACAGGCTGTGTGGATGAATTAAGTCG 204
OY	21 GlutleAlaArgValAspProLylyThrIlleglmetGylSerPhearglleAnaPro 40
Db	205 GAATTTGCCCGCGTGGACCCCAAGAACACATTCAGATGGGATCTCTCCGAATCAATCCA 264
OY	41 AspglySerGlnSerValValGluValProtyrAlaArgSerGluAlaHisLeuThrGlu 60
Db	265 GATGCACCACAATCGATTGTGGAGGTACTTATGCCCTCAGAGGCCACCTCCACAAG 324
OY	61 LeuleugluGluValAcysAspArgMetcytsGluTyrglygluGlnIlleAspProserThr 80
Db	325 TTGCTTGAGAGAGCTGTGTGACCGAATGAAGAGTAGTCGGGGACAAGATTGACCTTTCTACC 384
OY	81 HisArgLyAsnTyrrValArgValValSerArgAnsglyIuSerSerGluDeuAspLeu 100



•  
•  
•  
•

10

11

12



```

Db 165 GAATTCGCCCGGTGGACCCCAAGAACCATTCAGATGGGATCCTCCGAATCAATCCA 224
Oy 41 AspGlySerGlnSerValValGluValProTyrAlaArgSerGluAlaHisLeuThrGlu 60
Db 225 GATGCACGACGACTGAGTGTGGAGGTACTTATGCCCGCTCAGAGGCCACCTCAGAGAG 284
Oy 61 LeuLeuGluGluIleCysAspArgMetLysGluTyrGlyGluGlnIleAspProSerThr 80
Db 285 TTGCTTGAGGAGGTGTGTGACCGAATGAAGAGTAGGGGA-CAGATTGA-CCTTCTACC 342
Oy 81 HisArgLysAsnTyrValArgValValGlyArgAsnGlyGluSerSerGluLeuAspLeu 100
Db 343 CACCCCAAGAACTACGTACGCGCTCGTAGACCGGAATGGAAATCCAGTGAAGACTTA 402
Oy 101 GlnGlyIleArgIleAspSerAspIleSerGlyThrLeuLysPheAlaCysGluSerIle 120
Db 403 CAGGCGATCCGAAATTGACTCAGATATCANCGGCACCCCTCAAGTTGCCGTGAGAGCANT 462
Oy 121 ValGluGluTyrGluAspGluLeuIleGluPhePheSerArgGluAlaAspAsnValLys 140
Db 463 GTGGAAAGATACGAGATGAGCTTATCGAATCTTCTCCAGAGAGGTGACAACTTAA 522
Oy 141 AspLysLeuCysSerLysArgThrAspLeuCysAspHisAlaLeuHisIleSer-HisAs 160
Db 523 GACAAACTTTGACAGTAAAGCGGACAGACTTAATGTGACCAATCCCTGCACAGATCTCAGCA 582
Oy 160 pGluLeu 162
Db 583 AGAGCTC 589

```

Search completed: March 30, 2003, 22:42:58  
 Job time : 130.151 secs